

Bounding the Rate Region of Vector Gaussian Multiple Descriptions with Individual and Central Receivers

Guoqiang Zhang, W. Bastiaan Kleijn, *Fellow, IEEE*,
and Jan Østergaard, *Member, IEEE*

Abstract

In this work, the rate region of the vector Gaussian multiple description problem with individual and central quadratic distortion constraints is studied. In particular, an outer bound to the rate region of the L-description problem is derived. The bound is obtained by lower bounding a weighted sum rate for each supporting hyperplane of the rate region. The key idea is to introduce at most L-1 auxiliary random variables and further impose upon the variables a Markov structure according to the ordering of the description weights. This makes it possible to greatly simplify the derivation of the outer bound. In the scalar Gaussian case, the complete rate region is fully characterized by showing that the outer bound is tight. In this case, the optimal weighted sum rate for each supporting hyperplane is obtained by solving a single maximization problem. This contrasts with existing results, which require solving a min-max optimization problem.

Index Terms

Multiple description coding, rate region, entropy power inequality.

The work was partly presented at the Data Compression Conference, 2010.

G. Zhang (guoqiang.zhang@ee.kth.se) and W. B. Kleijn (bastiaan.kleijn@ee.kth.se) are with the school of Electrical Engineering, KTH-Royal Institute of Technology.

J. Østergaard (janoe@ieee.org) is with the Department of Electronic Systems, Aalborg University, Aalborg, Denmark.

I. INTRODUCTION

Multiple description (MD) coding is a joint source-channel coding scheme aimed at combating unreliable communication links. The basic principle is to generate a set of descriptions for the information source with the property that any subset of the descriptions provides an approximation of the source with a certain fidelity. A natural research goal is to determine the transmission limits under some quality of service requirement. Generally, there are $2^L - 1$ distortion constraints for the L -description case, each corresponding to a particular combination of received descriptions. Thus, the problem complexity increases exponentially with an increasing number of descriptions.

The pioneering work on the multiple description problem is the two-description achievable rate region by El Gamal and Cover [1] (EGC scheme). The region was shown to be tight for a memoryless Gaussian source by Ozarow [2], and was shown not to be tight in general when there is no excess rate (no redundancy) by Zhang and Berger [3]. This result was further extended to the L -description case in [4]. Later, the work in [5],[6], provided an enlarged achievable rate region by using the random binning ideas from distributed source coding.

The characterization of the (tight) rate region for the general L -description problem is quite challenging, and remains an open problem. Instead, the main focus has been on special cases, e.g., the case of symmetric side distortion constraints [7], or the case where only a subset of the distortion constraints is of concern [8], [9], [10].

For the particular case that only individual (only one description is received) and central receivers are of importance, Wang and Viswanath [8] derived the optimal sum rate for the vector Gaussian source. In [10], Chen considered the same transmission scenario, and derived the rate region. However, his work was limited to the scalar Gaussian source. The direct extension of his work to the vector Gaussian source appears to be difficult if possible.

The present work considers the rate region of the vector Gaussian multiple description problem with individual and central distortion constraints. One major contribution is that an outer bound is derived for the considered rate region. The outer bound is formulated by lower-bounding a weighted sum rate associated with a supporting hyper-plane of the rate region. The expression

for the bound only involves a maximization process.

Another contribution is that when the new approach is applied to the scalar Gaussian multiple description problem, the derived outer bound is shown to be tight, thus fully characterizing the rate region. Note that the expression for the bound corresponds to a maximization problem as compared to that in [10] which involves a min-max game (or a max-min game), thus significantly reducing the problem complexity.

We now summarize the notations used in this paper. Lower case letters denote scalar random variables and boldface lower case letters denote vector random variables. Boldface upper case letters are used to denote matrices. Specifically, we use $\mathbf{0}$ and \mathbf{I} to denote the all-zero matrices (including zero vectors) and the identity matrices, respectively. We also use \mathbf{H} to denote all-one matrices. The operator $|\cdot|$ refers to the determinant of a matrix, unless otherwise specified. The operator $\mathbb{E}(\cdot)$ denotes the expectation. For random vectors \mathbf{y}_1 and \mathbf{y}_2 , we use $\mathbb{E}[\mathbf{y}_1|\mathbf{y}_2]$ to denote the conditional expectation of \mathbf{y}_1 given \mathbf{y}_2 , and $\text{Cov}[\mathbf{y}_1|\mathbf{y}_2]$ to denote the covariance matrix of $\mathbf{y}_1 - \mathbb{E}[\mathbf{y}_1|\mathbf{y}_2]$. The partial order \succ (\succeq) denotes positive definite (semidefinite) ordering: $\mathbf{A} \succ \mathbf{B}$ ($\mathbf{A} \succeq \mathbf{B}$) means that $\mathbf{A} - \mathbf{B}$ is a positive definite (semidefinite) matrix. All the logarithm functions are to base e .

II. PROBLEM FORMULATION AND MAIN RESULTS

In this section we first define the multiple description rate-region problem formally. We then introduce the Gaussian multiple description scheme for the problem, which provides an achievable rate region. Finally we present the main results of the paper.

A. Problem Formulation

Suppose the information source is an i.i.d. process $\{\mathbf{x}[m]\}$ with marginal distribution $\mathcal{N}(\mathbf{0}, \mathbf{K}_x)$, i.e., a collection of i.i.d. real Gaussian random vectors. Let the covariance matrix \mathbf{K}_x be an $N \times N$ positive definite matrix. There are L encoding functions at the transmitter, each encoding a source sequence of length n , $\mathbf{x}^n = (\mathbf{x}[1]^t, \dots, \mathbf{x}[n]^t)^t$. Denote the resulting codewords as $f_l^{(n)}(\mathbf{x}^n)$, $l = 1, \dots, L$. Let $C_l^{(n)}$, $l = 1, \dots, L$, denote the corresponding codebooks, i.e.,

$f_l^{(n)}(x^n) \in C_l^{(n)}$. The output of the l 'th encoder is sent through the l 'th communication channel at rate $R_l = \frac{1}{n} \log |C_l^{(n)}|$, where $|C_l^{(n)}|$ denotes the cardinality of the codebook.

In response to the L descriptions, there are L individual receivers and one central receiver. The l 'th individual receiver uses its received codeword to generate an approximation $g_l(f_l^{(n)}(x^n))$ to the source sequence x^n , $l = 1, \dots, L$. On the other hand, the central receiver generates an approximation to the source sequence based on all L codewords. Since we are only interested in the quadratic distortion measure, the optimal approximation is given by the minimal mean-squared error (MMSE) estimation of the source sequence. We say a rate vector (R_1, \dots, R_L) is achievable if there exist, for all sufficiently large n , encoders of rates (R_1, \dots, R_L) and decoders, such that

$$\frac{1}{n} \sum_{m=1}^n \text{Cov} [x[m] | f_l^{(n)}(x^n)] \preceq D_l, \quad l = 1, \dots, L, \quad (1)$$

$$\frac{1}{n} \sum_{m=1}^n \text{Cov} [x[m] | f_1^{(n)}(x^n), \dots, f_L^{(n)}(x^n)] \preceq D_0. \quad (2)$$

The rate region $\mathcal{R}(K_x, D_1, \dots, D_L, D_0)$ is the convex hull of the set of all achievable rate vectors subject to the individual side distortion constraints D_l , $l = 1, \dots, L$, and the central distortion constraint D_0 . Throughout the paper, we consider $0 \prec D_0 \prec D_l$, $\forall l \in \{1, \dots, L\}$.

We focus on characterizing the rate region $\mathcal{R}(K_x, D_1, \dots, D_L, D_0)$. Our work is an extension of [8] which only addressed the optimal sum-rate problem for the above transmission scenario. There are two motivations behind our work. First obtaining the rate region is of great interest from a theoretic point of view. Second the rate-region expression can provide insight in designing efficient asymmetric (i.e, the rates are different across the channels) multiple description systems.

Motivated by the fact that $\mathcal{R}(K_x, D_1, \dots, D_L, D_0)$ is a closed convex set, we consider characterizing its supporting hyperplane, which can be formulated as an optimization problem of the form:

$$\min_{(R_1, \dots, R_L) \in \mathcal{R}(K_x, D_1, \dots, D_L, D_0)} \sum_{l=1}^L \beta_l R_l, \quad (3)$$

where β_l , $l = 1, \dots, L$, are arbitrary positive parameters. The case where at least one of the weighting factors is zero defines a trivial problem. To see this, notice that one can always put enough rate to the descriptions associated with the zero weighting factors, in order to meet the central distortion constraint. Then one simply needs to use the minimal single-description rate

for each remaining description in order to satisfy the individual distortion constraints and thereby achieve the minimal (optimal) weighted sum rate.

Without loss of generality, we may assume $\beta_1 \geq \dots \geq \beta_L > 0$, which can always be obtained by rearranging the description indices. It may happen that some weighting factors are equal. We further group the weighting factors $(\beta_2, \beta_2, \beta_3, \dots, \beta_L)$ (the first element is modified on purpose, and it will be explained later) according to their values. The elements within each group have the same value and different groups take different values. Denote the resulting weighting factors as $\alpha_1 > \alpha_2 > \dots > \alpha_J$, where α_j is associated with m_j rates in (3). With this, it follows that $\sum_{j=1}^J m_j = L$, and $m_1 \geq 2$. The maximum achievable value of J is $L - 1$. Denote

$$M_1^j = \sum_{i=1}^j m_i \quad j = 1, \dots, J \quad (4)$$

and let $M_1^0 = 0$. The optimization problem (3) can be rewritten as

$$\min_{(R_1, \dots, R_L) \in \mathcal{R}(\mathbf{K}_x, \mathbf{D}_1, \dots, \mathbf{D}_L, \mathbf{D}_0)} \alpha_0 R_1 + \sum_{j=1}^J \alpha_j \sum_{i=1}^{m_j} (R_{M_1^{j-1}+i}), \quad (5)$$

where $\alpha_0 = \beta_1 - \beta_2$. The reason to distinguish different weighting factors is to facilitate the derivation of the outer bound to (3). It will be shown later that the choice of α_0 is of little importance. It does not affect the construction of the outer bound.

B. Gaussian Description Scheme

In this subsection, we introduce the Gaussian description scheme, with which an achievable rate region is then obtained. Conceptually, in the transmission of a vector Gaussian source \mathbf{x} , L parallel noisy sub-channels with additive Gaussian noises are constructed. This has the advantage that the transmission distortions and rates can be easily analyzed. Let $\mathbf{w}_1, \dots, \mathbf{w}_L$ be N -dimensional Gaussian vectors independent of \mathbf{x} , of which the marginal distributions are denoted as $\mathcal{N}(\mathbf{0}, \mathbf{K}_l)$, $l = 1, \dots, L$. Define

$$\mathbf{u}_l = \mathbf{x} + \mathbf{w}_l, \quad l = 1, \dots, L. \quad (6)$$

Such a channel is referred to as a *Gaussian test channel* [11]. For any subset $S \subseteq \{1, \dots, L\}$, we use \mathbf{K}_S to denote the covariance matrix of all \mathbf{w}_l , $l \in S$. With a slight abuse of notation, we refer to $\mathbf{K}_{\{1, \dots, L\}}$ as \mathbf{K}_w .

We consider the MMSE estimation of \mathbf{x} using subsets of \mathbf{u}_l , $l = 1, \dots, L$. It can be shown that

$$\text{Cov}[\mathbf{x}|\mathbf{u}_l, l \in S] = (\mathbf{K}_x^{-1} + (\mathbf{I}_N, \dots, \mathbf{I}_N)\mathbf{K}_S^{-1}(\mathbf{I}_N, \dots, \mathbf{I}_N)^t)^{-1}, \quad \forall S \subseteq \{1, \dots, L\}. \quad (7)$$

Define

$$\underline{\mathbf{K}}_S \triangleq [(\mathbf{I}_N, \dots, \mathbf{I}_N)\mathbf{K}_S^{-1}(\mathbf{I}_N, \dots, \mathbf{I}_N)^t]^{-1}, \quad S \subseteq \{1, \dots, L\}. \quad (8)$$

It is immediate that $\underline{\mathbf{K}}_l = \mathbf{K}_l$, $l = 1, \dots, L$. Similarly to the introduction of \mathbf{K}_w , we let $\underline{\mathbf{K}}_w = \underline{\mathbf{K}}_{\{1, \dots, L\}}$. To clarify, for each $S \subseteq \{1, \dots, L\}$ with cardinality $|S|$, \mathbf{K}_S is of size $N|S| \times N|S|$, whereas $\underline{\mathbf{K}}_S$ is of size $N \times N$. Plugging (8) into (7) produces

$$\text{Cov}[\mathbf{x}|\mathbf{u}_l, l \in S] = (\mathbf{K}_x^{-1} + \underline{\mathbf{K}}_S^{-1})^{-1}, \quad S \subseteq \{1, \dots, L\}. \quad (9)$$

We can define a virtual additive Gaussian noisy channel for each subset S in (9), where the covariance matrix of the Gaussian noise takes the form $\underline{\mathbf{K}}_S$. Therefore, the quadratic distortion of the MMSE estimation of the source \mathbf{x} from this channel output gives the same expression as (9).

The achievable rate vector by using the Gaussian description scheme has been studied in detail in [4] (for the scalar source case), and in [8] (for the vector source case). We summarize the result in the following lemma.

Lemma 2.1 ([8]): For every \mathbf{K}_w such that

$$\begin{aligned} \text{Cov}[\mathbf{x}|\mathbf{u}_l] &\preceq \mathbf{D}_l, \quad l = 1, \dots, L \\ \text{Cov}[\mathbf{x}|\mathbf{u}_1, \dots, \mathbf{u}_L] &\preceq \mathbf{D}_0, \end{aligned} \quad (10)$$

the rate vector satisfying

$$\sum_{l \in S} R_l \geq \left[\sum_{l \in S} h(\mathbf{u}_l) \right] - h(\mathbf{u}_l, l \in S|\mathbf{x}) = \frac{1}{2} \log \frac{\prod_{l \in S} |\mathbf{K}_x + \mathbf{K}_l|}{|\mathbf{K}_S|}, \quad \forall S \subseteq \{1, \dots, L\}, \quad (11)$$

is achievable.

Lemma 2.1 defines an achievable rate region. In Section V, we will show that the achievable region is tight for a scalar Gaussian source.

C. Outer Bound to the Rate Region

In this subsection we present our new outer bound to the rate region $\mathcal{R}(\mathbf{K}_x, \mathbf{D}_1, \dots, \mathbf{D}_L, \mathbf{D}_0)$. The expression of the outer bound is obtained by lower-bounding a weighted sum rate (5) of every supporting hyper-plane of the rate region.

In our construction of the outer bound in our work, we introduce a set of auxiliary Gaussian random vectors as in [10]. Differently from [10], we impose a Markov structure on those auxiliary random vectors according to the order of the weighting factors. The main reason to group the weighting factors $(\beta_1, \beta_2, \beta_3, \dots, \beta_L)$ (see Subsection II-A) by their values is to identify the number of auxiliary Gaussian random vectors needed for the construction of the outer bound. For the optimization problem in (5), we introduce J auxiliary Gaussian random vectors, \mathbf{z}_j , $j = 1, \dots, J$, one for each distinct weighting factor (i.e., \mathbf{z}_j is associated with α_j). When all the weighting factors are identical (i.e., $J = 1$), only one auxiliary random vector is needed. This special case is for deriving an outer bound to the sum-rate, which was addressed in [8].

We now explain the Markov structure of the J random vectors in more detail. The covariance matrices of the auxiliary vectors are chosen such that a high-indexed vector has a large covariance matrix. Denote the covariance matrix of \mathbf{z}_j as \mathbf{N}_j , $j = 1, \dots, J$. The Markov structure can be mathematically formulated as $\mathbf{0} \prec \alpha_1 \mathbf{N}_1 \prec \alpha_2 \mathbf{N}_2 \prec \dots \prec \alpha_J \mathbf{N}_J$. By using this partial ordering structure, the derivation of the outer bound is significantly simplified. Further, it facilitates the recognition of optimality conditions for the outer bound to be tight. On the other hand, the work in [10] uses L auxiliary random variables without imposing a Markov structure on them.

Considering the multiple description coding with respect to individual and central distortion constraints for a Gaussian source with distribution $\mathcal{N}(\mathbf{0}, \mathbf{K}_x)$, we have the following outer bound to any supporting hyperplane of the rate region.

Theorem 2.2: Given the distortion constraints $(\mathbf{D}_1, \dots, \mathbf{D}_L, \mathbf{D}_0)$ and a zero-mean vector

Gaussian source with covariance matrix \mathbf{K}_x , the weighted sum rate (5) is lower-bounded by

$$\begin{aligned}
& \alpha_0 R_1 + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} R_{M_1^{j-1}+i} \right) \\
& \geq \sup_{B(\{\mathbf{N}_i\}_{i=1}^J)} \left(\frac{\alpha_0}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_1|} + \frac{\alpha_1}{2} \log \frac{\alpha_1^N |\mathbf{K}_x| |\mathbf{K}_x + \mathbf{N}_1|^{m_1-1} |\mathbf{N}_2 - \mathbf{N}_1|}{(\alpha_1 - \alpha_2)^N \prod_{i=1}^{m_1} |\mathbf{N}_1 + \mathbf{D}_i| |\mathbf{N}_2|} \right. \\
& \quad + \sum_{j=2}^{J-1} \frac{\alpha_j}{2} \log \frac{(\alpha_{j-1} - \alpha_j)^N |\mathbf{K}_x + \mathbf{N}_j|^{m_j} |\mathbf{N}_{j-1}| |\mathbf{N}_{j+1} - \mathbf{N}_j|}{(\alpha_j - \alpha_{j+1})^N \prod_{i=1}^{m_j} |\mathbf{N}_j + \mathbf{D}_{M_1^{j-1}+i}| |\mathbf{N}_j - \mathbf{N}_{j-1}| |\mathbf{N}_{j+1}|} \\
& \quad \left. + \frac{\alpha_J}{2} \log \frac{(\alpha_{J-1} - \alpha_J)^N |\mathbf{N}_J + \mathbf{D}_0| |\mathbf{K}_x + \mathbf{N}_J|^{m_J} |\mathbf{N}_{J-1}|}{\alpha_J^N |\mathbf{D}_0| \prod_{i=1}^{m_J} |\mathbf{N}_J + \mathbf{D}_{M_1^{J-1}+i}| |\mathbf{N}_J - \mathbf{N}_{J-1}|} \right), \tag{12}
\end{aligned}$$

where $B(\{\mathbf{N}_i\}_{i=1}^J) = \{(\mathbf{N}_1, \dots, \mathbf{N}_J) \in \mathbb{R}^{N \times NJ} | \mathbf{0} \prec \alpha_1 \mathbf{N}_1 \prec \alpha_2 \mathbf{N}_2 \prec \dots \prec \alpha_J \mathbf{N}_J\}$, and $J > 1$. For the special case that $J = 1$, the lower bound is expressed as

$$\alpha_0 R_1 + \alpha_1 \sum_{l=1}^L R_l \geq \sup_{\mathbf{N}_1 \succ \mathbf{0}} \frac{\alpha_0}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_1|} + \frac{\alpha_1}{2} \log \left(\frac{|\mathbf{K}_x| |\mathbf{K}_x + \mathbf{N}_1|^{L-1} |\mathbf{D}_0 + \mathbf{N}_1|}{|\mathbf{D}_0| \prod_{l=1}^L |\mathbf{D}_l + \mathbf{N}_1|} \right). \tag{13}$$

Proof: See Section III. ■

One special case of $B(\{\mathbf{N}_i\}_{i=1}^J)$ in (12) is of particular interest. Letting $\mathbf{N}_1 = [\epsilon/\alpha_1] \mathbf{I}_N$, $\mathbf{N}_j = [\epsilon(1 + \epsilon + \dots + \epsilon^{j-1})/\alpha_j] \mathbf{I}_N$, $j = 2, \dots, J$, and $\epsilon \rightarrow 0^+$, we obtain from (12) the following lower bound

$$\alpha_0 R_1 + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} R_{M_1^{j-1}+i} \right) \geq \frac{\alpha_0}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_1|} + \sum_{j=1}^J \frac{\alpha_j}{2} \left(\sum_{i=1}^{m_j} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_{M_1^{j-1}+i}|} \right). \tag{14}$$

This bound is actually the weighted summation of the L minimum single-description rates, $\frac{1}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_i|}$, $i = 1, \dots, L$. Note that the central distortion constraint is not involved in the bound. For the extreme case that the central distortion constraint is loose, (14) becomes a tight bound. In this situation, any hyperplane attains the rate region at the corner point $\left(\frac{1}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_1|}, \dots, \frac{1}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_L|} \right)$.

D. Optimal Weighted Sum Rate

In this subsection, we first present the optimality conditions for the Gaussian description scheme to produce the optimal weighted sum rate. Specifically, we derive the optimality conditions on the covariance matrix \mathbf{K}_w as introduced in subsection II-B. Then we discuss

how to compute the optimal matrix \mathbf{K}_w if it exists.

For the application of the Gaussian description scheme, we provide optimality conditions under which the performance of the scheme reaches the outer bound (12). Therefore, if a Gaussian description scheme can be constructed for a weighting vector $(\alpha_0, \alpha_1, \dots, \alpha_J)$ such that the optimality conditions are satisfied, the corresponding optimal weighted sum rate can be obtained.

Theorem 2.3: If, for any given \mathbf{K}_x , there exists a \mathbf{K}_w such that:

1) Layered correlation:

$$\mathbb{E} \left[\mathbf{w}_{M_1^{j-1}+i} \mathbf{w}_k^t \right] = -\mathbf{A}_j, \quad \forall 1 \leq k < M_1^{j-1} + i, i = 1, \dots, m_j, j = 1, \dots, J \quad (15)$$

2) Proportionality:

$$\alpha_j \left(\mathbf{A}_j + \underline{\mathbf{K}}_{\{1, \dots, M_1^j\}} \right)^{-1} - (\alpha_j - \alpha_{j+1}) \underline{\mathbf{K}}_{\{1, \dots, M_1^j\}}^{-1} = \alpha_{j+1} \left(\mathbf{A}_{j+1} + \underline{\mathbf{K}}_{\{1, \dots, M_1^j\}} \right)^{-1}, \quad j = 1, \dots, J-1, \quad (16)$$

where $\mathbf{0} \prec \alpha_1 \mathbf{A}_1 \prec \alpha_2 \mathbf{A}_2 \prec \dots \prec \alpha_J \mathbf{A}_J \prec \alpha_J \mathbf{K}_x$, and such that the set of distortion constraints $(\mathbf{D}_1, \dots, \mathbf{D}_L, \mathbf{D}_0)$ is achieved with equality, then the outer bound given by (12) is tight.

Proof: See Section IV. ■

The property of layered correlation in Theorem 2.3 refers to the fact that the correlation matrix of \mathbf{w}_j and any lower indexed random vector \mathbf{w}_k $k < j$ remains the same. The number of different correlation matrices is determined by the number of different weighting factors. For the optimization problem (5), J different correlation matrices have to be constructed. Informally, each correlation matrix \mathbf{A}_j in (15) controls the redundancy among the m_j descriptions associated with α_j , and the set of descriptions associated with $\{\alpha_i, i < j\}$. Further, a small correlation matrix corresponds to a high redundancy among the associated descriptions. For the extreme case that the central distortion constraint is loose, the optimal correlation matrices take the form $\mathbf{A}_1 = \dots = \mathbf{A}_J = \mathbf{0}$. The L descriptions are highly correlated, implying a high redundancy embedded in the descriptions. The optimal weighted sum rate for this special case is given by (14).

The proportionality condition (16) imposes an ordered structure on the correlation matrices,

i.e., $\mathbf{0} \prec \alpha_1 \mathbf{A}_1 \prec \alpha_2 \mathbf{A}_2 \prec \dots \prec \alpha_J \mathbf{A}_J \prec \alpha_J \mathbf{K}_x$. This ordered structure is essentially determined by the ordered structure on the weighting factors, i.e., $\alpha_1 > \alpha_2 > \dots > \alpha_J$. We now explain this property in an informal way. If the central distortion constraint is active, each description has to carry an extra rate besides the minimum single-description rate to account for the central distortion constraint. Since a large weighting factor enforces a large penalty on the associated description rate in (5), it is desirable that the descriptions with a large weighting factor carry a small amount of extra rates. By relating a small correlation matrix with a large weighting factor (i.e., \mathbf{A}_j corresponds to α_j , $j = 1, \dots, J$), the redundancy introduced by the correlation matrix is large, rendering a small amount of extra rate for the descriptions associated with the large weighting factor. (16) fully characterizes the relationship between the correlation matrices and the weighting factors.

Next we consider how to obtain the the covariance matrix \mathbf{K}_w satisfying the conditions in Theorem 2.3 if it exists. We first study the property of the layered correlation (15). We find that if the covariance matrix \mathbf{K}_w has the layered structure (15), the matrices $\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}}$ and \mathbf{A}_j , $j = 1, \dots, J$ carry some simple relationships. We present the result in the following lemma.

Lemma 2.4: Suppose the covariance matrix \mathbf{K}_w satisfies (15). If $\mathbf{K}_w \succ \mathbf{0}$ and $\mathbf{A}_j \succeq \mathbf{0}$, $j = 1, \dots, J$, then

$$\begin{cases} \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^1\}} + \mathbf{A}_1 \right)^{-1} = \sum_{i=1}^{m_1} (\mathbf{K}_i + \mathbf{A}_1)^{-1} \\ \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}} + \mathbf{A}_j \right)^{-1} = \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{A}_j \right)^{-1} + \sum_{i=1}^{m_j} \left(\mathbf{K}_{M_1^{j-1}+i} + \mathbf{A}_j \right)^{-1}, \quad j = 2, \dots, J \end{cases} \quad (17)$$

Proof: See Appendix C. ■

Equ. (16) together with (17) fully characterizes the relationships among $\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}}$ and \mathbf{A}_j , $j = 1, \dots, J$. Note that given $\{\mathbf{K}_l : l = 1, \dots, L\}$, (16)-(17) can be interpreted as a set of functions \mathbf{A}_j , $j = 2, \dots, J$, and $\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}}$, $j = 1, \dots, J$ over \mathbf{A}_1 . In other words, \mathbf{A}_1 is the only free covariance matrix under (16)-(17). The problem for obtaining \mathbf{K}_w is now reduced to determine the correlation matrix \mathbf{A}_1 and $\{\mathbf{K}_l : l = 1, \dots, L\}$. In fact, the covariance matrices $\{\mathbf{K}_l : l = 1, \dots, L\}$ can be determined by the individual side distortion constraints. On the other hand, the correlation matrix \mathbf{A}_1 can be determined by the central distortion constraint.

We now revisit the side and central distortion constraints. Suppose the distortion constraints $(\mathbf{D}_1, \dots, \mathbf{D}_L, \mathbf{D}_0)$ are achieved with equality using the Gaussian description scheme. By applying (9), we have

$$\begin{aligned}\mathbf{K}_l &= (\mathbf{D}_l^{-1} - \mathbf{K}_x^{-1})^{-1}, \quad l = 1, \dots, L \\ \underline{\mathbf{K}}_w &= (\mathbf{D}_0^{-1} - \mathbf{K}_x^{-1})^{-1}.\end{aligned}\tag{18}$$

If there exists a positive definite solution \mathbf{A}_1 to (16)-(18), then the distortion constraints are met with equality. The remaining work is to check if the matrix \mathbf{K}_w constructed from the solutions of (16)-(18) is positive definite. It turns out that as long as \mathbf{A}_1 is a solution to (16)-(17) where $\{\mathbf{K}_l \succ \mathbf{0} : l = 1, \dots, L\}$ and $\underline{\mathbf{K}}_w \succ \mathbf{0}$ are constant matrices, the resulting \mathbf{K}_w is always positive definite. We state this formally below.

Lemma 2.5: Let $\underline{\mathbf{K}}_w \succ \mathbf{0}$ and $\mathbf{K}_l \succ \mathbf{0}$, $l = 1, \dots, L$ be constant matrices in (16)-(17). If for some $\mathbf{A}_1 \succ \mathbf{0}$, (16)-(17) are true and $\mathbf{0} \prec \alpha_1 \mathbf{A}_1 \prec \alpha_2 \mathbf{A}_2 \prec \dots \prec \alpha_J \mathbf{A}_J$, then the matrix \mathbf{K}_w constructed using (15) is positive definite.

Proof: See Appendix E. ■

We summarize the result in the following proposition.

Proposition 2.6: Given distortion constraints $(\mathbf{D}_1, \dots, \mathbf{D}_L, \mathbf{D}_0)$. If there exists a solution \mathbf{A}_1^* to (16)-(18), then the Gaussian description scheme with \mathbf{K}_w constructed using (15) achieves the optimal weighted sum rate. The optimal \mathbf{N}_j in (12) is given as $\mathbf{N}_j = (\mathbf{A}_j(\mathbf{A}_1^*)^{-1} - \mathbf{K}_x^{-1})^{-1}$, $j = 1, \dots, J$.

E. Rate Region for the Scalar Gaussian Source

In this subsection, we use Theorem 2.3 to prove that for the scalar Gaussian case, the outer bound on the weighted sum-rate provided by (12) is tight. This completely characterizes the rate region for the scalar Gaussian case and parallels a recent result of Chen [10].

Theorem 2.7: For a scalar Gaussian source and individual and central distortion constraints, the complete rate region is given by (12) and (13).

Proof: See Section V. ■

Remark 2.8: It should be noted from Theorem 2.7 and (12)-(13), that the scalar Gaussian rate region is obtained by solving a single maximization problem, whereas the approach of [10] requires the solution of a min-max optimization problem.

III. PROOF OF THEOREM 2.2

Before presenting the proof of Theorem 2.2, we need the following lemma:

Lemma 3.1: Assume that \mathbf{v} and \mathbf{x}^n are arbitrarily distributed random variables (which might be correlated) and assume that given \mathbf{v} , \mathbf{x}^n has a density. Denote the dimensionality of \mathbf{x}^n as nN . Let $\{\mathbf{z}_1[m]\}_{m=1}^n$ and $\{\mathbf{z}_2[m]\}_{m=1}^n$ be two i.i.d. random Gaussian vector processes with marginal distributions $\mathcal{N}(\mathbf{0}, \mathbf{N}_1)$ and $\mathcal{N}(\mathbf{0}, \mathbf{N}_2)$, respectively. Both \mathbf{N}_1 and \mathbf{N}_2 are of size $N \times N$. Let $\mathbf{z}_1^n = [\mathbf{z}_1[1]^t, \dots, \mathbf{z}_1[n]^t]^t$. In a similar manner, we can define \mathbf{z}_2^n . Both \mathbf{z}_1^n and \mathbf{z}_2^n are independent of \mathbf{v} and \mathbf{x}^n . Let $\mu_1 > \mu_2 > 0$ and $\mathbf{0} \prec \mu_1 \mathbf{N}_1 \prec \mu_2 \mathbf{N}_2$. Then there is

$$\begin{aligned} & \mu_2 h(\mathbf{x}^n + \mathbf{z}_2^n | \mathbf{v}) - \mu_1 h(\mathbf{x}^n + \mathbf{z}_1^n | \mathbf{v}) + (\mu_1 - \mu_2) h(\mathbf{x}^n | \mathbf{v}) \\ & \leq \frac{n\mu_1}{2} \log \frac{(\mu_1 - \mu_2)^N |\mathbf{N}_2|}{\mu_1^N |\mathbf{N}_2 - \mathbf{N}_1|} - \frac{n\mu_2}{2} \log \frac{(\mu_1 - \mu_2)^N |\mathbf{N}_1|}{\mu_2^N |\mathbf{N}_2 - \mathbf{N}_1|}, \end{aligned} \quad (19)$$

where the equality holds if \mathbf{x}^n and \mathbf{v} are jointly Gaussian and satisfy

$$\text{Cov}[\mathbf{x}^n | \mathbf{v}] = \mathbf{I}_n \otimes (\mu_1 - \mu_2) \mathbf{N}_2 (\mu_2 \mathbf{N}_2 - \mu_1 \mathbf{N}_1)^{-1} \mathbf{N}_1, \quad (20)$$

where \otimes denotes the Kronecker product [12].

Proof: See Appendix B. ■

When the problem specializes to the scalar source case (i.e. $N = 1$), an alternative upper bound of the right hand side of (19) was derived by Chen [10]. The main difference is that the upper bound in [10] requires solving an optimization problem, and does not admit a simple closed-form expression.

We now proceed to present the argument for the outer bound. We first consider the case $J > 1$. We define J i.i.d. random Gaussian processes $\mathbf{z}_i^n = \{\mathbf{z}_i[m]\}_{m=1}^n$ with marginal distributions $\mathcal{N}(\mathbf{0}, \mathbf{N}_j)$, $j = 1, \dots, J$, respectively. Let $C_l = f_l^{(n)}(\mathbf{x}^n)$, $l = 1, \dots, L$, be the discrete random variables. Suppose that $\mathbf{z}_1^n, \dots, \mathbf{z}_J^n$ are independent of \mathbf{x}^n and C_l , $l = 1, \dots, L$. Next we introduce

J processes $\mathbf{y}_j^n = (\mathbf{y}_j[1]^t, \dots, \mathbf{y}_j[n]^t)^t$, $j = 1, \dots, J$ by

$$\mathbf{y}_j[m] = \mathbf{x}[m] + \mathbf{z}_j[m], \quad m = 1, \dots, n \text{ and } j = 1, \dots, J. \quad (21)$$

It follows that each sequence $\{\mathbf{y}_j[m]\}$ is an i.i.d. process with marginal distribution $\mathcal{N}(\mathbf{0}, \mathbf{K}_{y_j})$, where $\mathbf{K}_{y_j} = \mathbf{K}_x + \mathbf{N}_j$, $j = 1, \dots, J$. The following sequence of inequalities provides a lower bound to the weighted sum rate (5):

$$\begin{aligned} & n\alpha_0 R_1 + \sum_{j=1}^J n\alpha_j \left(\sum_{i=1}^{m_j} R_{M_1^{j-1}+i} \right) \\ & \geq \alpha_0 H(C_1) + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} H(C_{M_1^{j-1}+i}) \right) \\ & \stackrel{(a)}{\geq} \alpha_0 [H(C_1) - H(C_1|\mathbf{x}^n)] + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} H(C_{M_1^{j-1}+i}) - [H(C_1, \dots, C_{M_1^{j-1}}|\mathbf{y}_j^n) + \right. \\ & \quad \left. \sum_{i=1}^{m_j} H(C_{M_1^{j-1}+i}|\mathbf{y}_j^n) - H(C_1, \dots, C_{M_1^j}|\mathbf{y}_j^n)] \right) \\ & = \alpha_0 [h(\mathbf{x}^n) - h(\mathbf{x}^n|C_1)] + \sum_{j=1}^J \alpha_j \sum_{i=1}^{m_j} I(\mathbf{y}_j^n; C_{M_1^{j-1}+i}) + \alpha_J [H(C_1, \dots, C_L|\mathbf{y}_J^n) - H(C_1, \dots, C_L|\mathbf{x}^n)] + \\ & \quad \sum_{j=2}^J [\alpha_{j-1} H(C_1, \dots, C_{M_1^{j-1}}|\mathbf{y}_{j-1}^n) - \alpha_j H(C_1, \dots, C_{M_1^{j-1}}|\mathbf{y}_j^n) - (\alpha_{j-1} - \alpha_j) H(C_1, \dots, C_{M_1^{j-1}}|\mathbf{x}^n)] \\ & = \alpha_0 [h(\mathbf{x}^n) - h(\mathbf{x}^n|C_1)] + \sum_{j=1}^J \alpha_j \sum_{i=1}^{m_j} (h(\mathbf{y}_j^n) - h(\mathbf{y}_j^n|C_{M_1^{j-1}+i})) + \\ & \quad \sum_{j=2}^J [\alpha_{j-1} h(\mathbf{y}_{j-1}^n|C_1, \dots, C_{M_1^{j-1}}) - \alpha_j h(\mathbf{y}_j^n|C_1, \dots, C_{M_1^{j-1}}) - (\alpha_{j-1} - \alpha_j) h(\mathbf{x}^n|C_1, \dots, C_{M_1^{j-1}})] \\ & \quad + \alpha_J [h(\mathbf{y}_J^n|C_1, \dots, C_L) - h(\mathbf{x}^n|C_1, \dots, C_L)] + \alpha_1 h(\mathbf{x}^n) - \alpha_1 h(\mathbf{y}_1^n). \end{aligned} \quad (22)$$

The construction of the inequality (a) is crucial in the outer-bound derivation. For each distinct weighting factor, we essentially introduce an auxiliary Gaussian process.

We now consider deriving a lower bound to (22). As \mathbf{x}^n and \mathbf{y}_j^n , $j = 1, \dots, L$ are Gaussian vectors, we easily obtain

$$\begin{aligned} h(\mathbf{x}^n) &= \frac{1}{2} \log(2\pi e)^{Nn} |\mathbf{K}_x|^n, \\ h(\mathbf{y}_j^n) &= \frac{1}{2} \log(2\pi e)^{Nn} |\mathbf{K}_x + \mathbf{N}_j|^n, \quad j = 1, \dots, J. \end{aligned} \quad (23)$$

We will also use some entropy-related inequalities developed in [8], which are

$$h(\mathbf{x}^n|C_1) \leq \frac{1}{2} \log(2\pi e)^{Nn} |\mathbf{D}_1|^n, \quad (24)$$

$$h(\mathbf{y}_j^n|C_{M_1^{j-1}+i}) \leq \frac{1}{2} \log(2\pi e)^{Nn} |\mathbf{D}_{M_1^{j-1}+i} + \mathbf{N}_j|^n, \quad i = 1, \dots, m_j \text{ and } j = 1, \dots, J, \quad (25)$$

$$h(\mathbf{y}_J^n|C_1, \dots, C_L) - h(\mathbf{x}^n|C_1, \dots, C_L) \geq \frac{n}{2} \log \frac{|\mathbf{D}_0 + \mathbf{N}_J|}{|\mathbf{D}_0|}. \quad (26)$$

By using Lemma 3.1, the remaining quantities in (22) can be lower-bounded as

$$\begin{aligned} & \alpha_{j-1} h(\mathbf{y}_{j-1}^n|C_1, \dots, C_{M_1^{j-1}}) - \alpha_j h(\mathbf{y}_j^n|C_1, \dots, C_{M_1^{j-1}}) - (\alpha_{j-1} - \alpha_j) h(\mathbf{x}^n|C_1, \dots, C_{M_1^{j-1}}) \\ & \geq -\frac{n\alpha_{j-1}}{2} \log \frac{(\alpha_{j-1} - \alpha_j)^N |\mathbf{N}_j|}{\alpha_{j-1}^N |\mathbf{N}_j - \mathbf{N}_{j-1}|} + \frac{n\alpha_j}{2} \log \frac{(\alpha_{j-1} - \alpha_j)^N |\mathbf{N}_{j-1}|}{\alpha_j^N |\mathbf{N}_j - \mathbf{N}_{j-1}|}, \quad j = 2, \dots, J, \end{aligned} \quad (27)$$

where $\mathbf{0} \prec \alpha_1 \mathbf{N}_1 \prec \dots \prec \alpha_J \mathbf{N}_J$. The equalities hold in (27) if C_1, \dots, C_L and \mathbf{x}^n are jointly Gaussian with conditional covariance matrices

$$\text{Cov}[\mathbf{x}^n|C_1, \dots, C_{M_1^{j-1}}] = \mathbf{I}_n \otimes (\alpha_{j-1} - \alpha_j) \mathbf{N}_j (\alpha_j \mathbf{N}_j - \alpha_{j-1} \mathbf{N}_{j-1})^{-1} \mathbf{N}_{j-1}, \quad j = 2, \dots, J. \quad (28)$$

Plugging (23)-(27) into (22) produces the expression (12).

For the case $J = 1$ (i.e., the sum-rate case), we can follow the same derivation steps as those for $J > 1$. The argument for this special case is actually the same as that in [8]. An alternative derivation is to start with the expression of the outer bound for $J > 1$, and let $\alpha_2, \dots, \alpha_J$ approach α_1 in the expression. The proof is complete.

IV. PROOF OF THEOREM 2.3

In this section we study the optimality of the Gaussian description scheme in achieving the outer bound (12). We identify the optimality conditions by investigating the inequalities used in the outer-bound derivation under the Gaussian description scheme.

Similarly to the derivation of the outer bound, we introduce J auxiliary Gaussian random vectors $\mathbf{z}_j \sim \mathcal{N}(\mathbf{0}, \mathbf{N}_j)$, $j = 1, \dots, J$, independent of \mathbf{x} and all \mathbf{w}_l 's. As for (27), we put a partial ordering constraint on \mathbf{N}_j , $l = 1, \dots, J$:

$$\mathbf{0} \prec \alpha_1 \mathbf{N}_1 \prec \dots \prec \alpha_J \mathbf{N}_J. \quad (29)$$

To simplify the discussion, we introduce $\alpha_{J+1} = 0$. Letting $\mathbf{y}_j = \mathbf{x} + \mathbf{z}_j$, $j = 1, \dots, J$, we

obtain the following lower bound to (5) under the Gaussian description scheme:

$$\begin{aligned}
& \alpha_0 R_1 + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} R_{M_1^{j-1}+i} \right) \\
&= \alpha_0 R_1 + \sum_{j=1}^J (\alpha_j - \alpha_{j+1}) \left(\sum_{i=1}^{M_1^j} R_i \right) \\
&\stackrel{(a)}{=} \alpha_0 (h(\mathbf{u}_1) - h(\mathbf{u}_1|\mathbf{x})) + \sum_{j=1}^J (\alpha_j - \alpha_{j+1}) \left(\sum_{i=1}^{M_1^j} h(\mathbf{u}_i) - h(\mathbf{u}_1, \dots, \mathbf{u}_{M_1^j}|\mathbf{x}) \right) \\
&= \alpha_0 I(\mathbf{u}_1; \mathbf{x}) - \sum_{j=1}^J (\alpha_j - \alpha_{j+1}) h(\mathbf{u}_1, \dots, \mathbf{u}_{M_1^j}|\mathbf{x}) + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} h(\mathbf{u}_{M_1^{j-1}+i}) \right) \\
&\stackrel{(b)}{\geq} \alpha_0 I(\mathbf{u}_1; \mathbf{x}) - \sum_{j=1}^J (\alpha_j - \alpha_{j+1}) h(\mathbf{u}_1, \dots, \mathbf{u}_{M_1^j}|\mathbf{x}) + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} h(\mathbf{u}_{M_1^{j-1}+i}) \right. \\
&\quad \left. - \left[h(\mathbf{u}_1, \dots, \mathbf{u}_{M_1^{j-1}}|\mathbf{y}_j) + \sum_{i=1}^{m_j} h(\mathbf{u}_{M_1^{j-1}+i}|\mathbf{y}_j) - h(\mathbf{u}_1, \dots, \mathbf{u}_{M_1^j}|\mathbf{y}_j) \right] \right) \\
&= \alpha_0 I(\mathbf{u}_1; \mathbf{x}) + \sum_{j=1}^J \alpha_j \left(\sum_{i=1}^{m_j} I(\mathbf{u}_{M_1^{j-1}+i}; \mathbf{y}_j) \right) + \alpha_J [h(\mathbf{y}_J|\mathbf{u}_1, \dots, \mathbf{u}_L) - h(\mathbf{x}|\mathbf{u}_1, \dots, \mathbf{u}_L)] \\
&\quad + \sum_{j=2}^J \left[\alpha_{j-1} h(\mathbf{y}_{j-1}|\mathbf{u}_1, \dots, \mathbf{u}_{M_1^{j-1}}) - \alpha_j h(\mathbf{y}_j|\mathbf{u}_1, \dots, \mathbf{u}_{M_1^{j-1}}) - (\alpha_{j-1} - \alpha_j) h(\mathbf{x}|\mathbf{u}_1, \dots, \mathbf{u}_{M_1^{j-1}}) \right] \\
&\quad + \alpha_1 h(\mathbf{x}) - \alpha_1 h(\mathbf{y}_1), \tag{30}
\end{aligned}$$

where step (a) follows from Lemma 2.1.

We now study the inequality (b) in the derivation of (30). It is straightforward that if

$$h(\mathbf{u}_1, \dots, \mathbf{u}_{M_1^{j-1}}|\mathbf{y}_j) + \sum_{i=1}^{m_j} h(\mathbf{u}_{M_1^{j-1}+i}|\mathbf{y}_j) - h(\mathbf{u}_1, \dots, \mathbf{u}_{M_1^j}|\mathbf{y}_j) = 0, \quad j = 1, \dots, J, \tag{31}$$

then (b) is actually an equality. The fact that (31) hold implies that conditioned on \mathbf{y}_j , $\mathbf{u}_{M_1^{j-1}+1}, \dots, \mathbf{u}_{M_1^j}$, and $[\mathbf{u}_1^t, \dots, \mathbf{u}_{M_1^{j-1}}^t]^t$ are independent. In other words, the m_j descriptions associated with α_j , and the set of descriptions associated with $\{\alpha_i, i < j\}$, are conditionally independent given \mathbf{y}_j . Further, the conditional independencies (31) exhibit a layered structure. The conditional independency w.r.t. \mathbf{y}_j only refers to those descriptions with weighting factors no larger than α_j . The number of layered conditional independency is determined by the number of distinct weighting factors. Note that all the random variables involved in (30) are Gaussian vectors. The condition (31) can be equivalently described by constraints on their covariance

matrices. We present these constraints in the following proposition.

Proposition 4.1: There exist a set of matrices $(\mathbf{N}_1, \dots, \mathbf{N}_J)$ such that (29) holds and (31) is true if the covariance matrix \mathbf{K}_w satisfies the following three conditions

- equation (15) is true;
- the covariance matrix \mathbf{N}_j takes the form

$$\mathbf{N}_j = (\mathbf{A}_j^{-1} - \mathbf{K}_x^{-1})^{-1}, \quad j = 1, \dots, J. \quad (32)$$

- The correlation matrices \mathbf{A}_j , $j = 1, \dots, J$ satisfy a partial ordering constraint:

$$\mathbf{0} \prec \alpha_1 \mathbf{A}_1 \prec \alpha_2 \mathbf{A}_2 \prec \dots \prec \alpha_J \mathbf{A}_J \prec \alpha_J \mathbf{K}_x. \quad (33)$$

Proof:

We first derive the conditions such that (31) is true. Then we consider the partial ordering constraint (29) to obtain the additional conditions.

Note that the J conditional independencies in (31) have the same structure. We can focus on one particular case. The argument for other cases are the same. For a particular j , we use similar derivation steps as in the proof of [8, Proposition 2] to obtain the conditions. We find that (31) holds for each j when

$$\mathbb{E} \left[\mathbf{w}_{M_1^{j-1}+i} \mathbf{w}_k \right] = -\mathbf{A}_j, \quad \forall 1 \leq k < M_1^{j-1} + i, \quad i = 1, \dots, m_j$$

$$\mathbf{0} \prec \mathbf{A}_j \prec \mathbf{K}_x$$

$$\mathbf{N}_j = (\mathbf{A}_j^{-1} - \mathbf{K}_x^{-1})^{-1}.$$

Proposition 2 in [8] actually considered the conditional independency for $J = 1$ (the sum-rate case).

Next we show that (32) and (33) together are sufficient to produce (29). By using Lemma A.2, we have

$$\begin{aligned} \alpha_j \mathbf{A}_j \prec \alpha_{j+1} \mathbf{A}_{j+1} &\Leftrightarrow \alpha_{j+1} \mathbf{A}_j^{-1} \succ \alpha_j \mathbf{A}_{j+1}^{-1} \\ &\Rightarrow \alpha_{j+1} \mathbf{A}_j^{-1} + (\alpha_j - \alpha_{j+1}) \mathbf{K}_x^{-1} \succ \alpha_j \mathbf{A}_{j+1}^{-1} \\ &\Leftrightarrow \alpha_{j+1} (\mathbf{A}_j^{-1} - \mathbf{K}_x^{-1}) \succ \alpha_j (\mathbf{A}_{j+1}^{-1} - \mathbf{K}_x^{-1}) \end{aligned}$$

$$\Leftrightarrow \alpha_j \mathbf{N}_j \prec \alpha_{j+1} \mathbf{N}_{j+1}, \quad j = 1, \dots, J-1.$$

The proof is complete. ■

We consider the matrix \mathbf{K}_w satisfying the conditions in Proposition 4.1. We proceed to recognize the additional conditions such that (30) reaches the outer bound (12). Again since all the random variables involved in (30) are Gaussian vectors, knowing their covariance matrices is sufficient to characterize the expression. We list the distortion conditions as

$$\text{Cov}[\mathbf{x}|\mathbf{u}_l] = \mathbf{D}_l, \quad l = 1, \dots, L, \quad (34)$$

$$\text{Cov}[\mathbf{x}|\mathbf{u}_1, \dots, \mathbf{u}_{M_1^{j-1}}] = (\mathbf{K}_x^{-1} + \frac{\alpha_j}{\alpha_{j-1} - \alpha_j} \mathbf{A}_{j-1}^{-1} - \frac{\alpha_{j-1}}{\alpha_{j-1} - \alpha_j} \mathbf{A}_j^{-1})^{-1}, \quad j = 2, \dots, J, \quad (35)$$

$$\text{Cov}[\mathbf{x}|\mathbf{u}_1, \dots, \mathbf{u}_L] = \mathbf{D}_0. \quad (36)$$

The above distortion conditions are derived from (23)-(28) and (32). In particular, (35) is obtained by combining (28) and (32). (34) and (36) are for the individual side distortion constraints and the central distortion constraint, respectively. They correspond to (24)-(26). By using (32) and (34)-(36), it can be shown that (30) gives the same expression as the outer bound (12).

It is now clear that (15), (33) and (34)-(36) are optimality conditions on the covariance matrix \mathbf{K}_w for the weighted sum rate. When a proper \mathbf{K}_w can be constructed satisfying the optimality conditions, the outer bound (12) is tight under the expression (32). From (9), the condition (35) can be rewritten as

$$\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}}^{-1} = \frac{\alpha_{j+1}}{\alpha_j - \alpha_{j+1}} \mathbf{A}_j^{-1} - \frac{\alpha_j}{\alpha_j - \alpha_{j+1}} \mathbf{A}_{j+1}^{-1}, \quad j = 1, \dots, J-1. \quad (37)$$

By using Lemma A.1 on (37), we then arrive at the *proportionality* condition (16). For each $j = 1, \dots, J-1$, the expressions (16) and (37) are equivalent when \mathbf{A}_j and \mathbf{A}_{j+1} are positive definite. However, the condition (16) is more powerful in that it can handle the case that $\mathbf{A}_j \succeq \mathbf{0}$, $j = 1, \dots, J$. The generality of (16) over (37) might be useful to extend Theorem 2.3 to the case that the central distortion is loose. In our work, we will not discuss this case. The proof is complete.

V. RATE REGION FOR SCALAR GAUSSIAN SOURCE

In this section we consider the rate region of the multiple description problem with individual and central receivers for a scalar Gaussian source $x \sim \mathcal{N}(0, \sigma_x^2)$. Denote the individual and central distortion constraints as d_l , $l = 1, \dots, L$, and d_0 , where $0 < d_0 < d_l < \sigma_x^2$ for all $l = 1, \dots, L$. We construct the Gaussian test channel with w_l , $l = 1, \dots, L$, such that

$$\begin{aligned} k_l &= \text{Cov}[w_l], \\ \mathbb{E}[w_{M_1^{j-1}+i} w_k] &= -\sigma_j^2 \quad \forall 1 \leq k < M_1^{j-1} + i, \quad i = 1, \dots, m_j, \text{ and } j = 1, \dots, J. \end{aligned} \quad (38)$$

The correlation coefficient σ_j^2 corresponds to \mathbf{A}_j in (15) defined for general vector case. With this Gaussian description scheme, we find that the outer bound in *Theorem 2.2* is tight.

It is evident from Proposition 2.6 that the conditions (16) and (17) play an important role in establishing optimality of a Gaussian test channel. Note that for the scalar case, (16) is equivalent to (37) when the correlation coefficients $\sigma_j^2 > 0$, $j = 1, \dots, J$. We study the properties of σ_j^2 , $j = 2, \dots, J$ and $\underline{k}_{\{1, \dots, M_1^j\}}$, $j = 1, \dots, J$ as functions of σ_1^2 through (37) and (17).

Lemma 5.1: Let $k_l > 0$, $l = 1, \dots, L$, be constants. Define σ_j^2 , $j = 2, \dots, J$, and $\underline{k}_{\{1, \dots, M_1^j\}}$, $j = 1, \dots, J$ as functions of σ_1^2 , expressed as

$$\begin{cases} \left(\underline{k}_{\{1, \dots, M_1^1\}} + \sigma_1^2 \right)^{-1} = \sum_{i=1}^{m_1} (\sigma_1^2 + k_i)^{-1} \\ \left(\underline{k}_{\{1, \dots, M_1^j\}} + \sigma_j^2 \right)^{-1} = \left(\underline{k}_{\{1, \dots, M_1^{j-1}\}} + \sigma_j^2 \right)^{-1} + \sum_{i=1}^{m_j} \left(\sigma_j^2 + k_{M_1^{j-1}+i} \right)^{-1}, \quad j = 2, \dots, J \end{cases}, \quad (39)$$

$$\text{and } \underline{k}_{\{1, \dots, M_1^j\}}^{-1} = \frac{\alpha_{j+1}}{(\alpha_j - \alpha_{j+1})\sigma_j^2} - \frac{\alpha_j}{(\alpha_j - \alpha_{j+1})\sigma_{j+1}^2}, \quad j = 1, \dots, J-1. \quad (40)$$

Then there exists $\ddot{\sigma}_1^2 > 0$ such that the variables σ_l^2 , $l = 2, \dots, L$, are monotonically increasing over $\sigma_1^2 \in (0, \ddot{\sigma}_1^2)$, and the variables $\underline{k}_{\{1, \dots, M_1^j\}}$, $j = 1, \dots, J$, are monotonically decreasing over $\sigma_1^2 \in (0, \ddot{\sigma}_1^2)$. For any $\sigma_1^2 \in (0, \ddot{\sigma}_1^2)$, there is

$$0 < \alpha_1 \sigma_1^2 < \alpha_2 \sigma_2^2 < \dots < \alpha_J \sigma_J^2. \quad (41)$$

Further,

$$\sigma_j^2 \rightarrow \sigma_{j-1}^2 \text{ as } \sigma_1^2 \rightarrow 0, \quad \text{where } j = 2, \dots, J \quad (42)$$

$$\underline{k}_w \rightarrow \left(\sum_{i=1}^L k_i^{-1} \right)^{-1} \text{ as } \sigma_1^2 \rightarrow 0, \quad (43)$$

$$\underline{k}_w \rightarrow 0 \text{ as } \sigma_1^2 \rightarrow \tilde{\sigma}_1^2. \quad (44)$$

Proof: See Appendix D for the proof. ■

Upon establishing the result in *Lemma 5.1*, we are now ready to argue that for the scalar Gaussian source, the multiple description scheme using Gaussian descriptions achieves the outer bound in *Theorem 2.2*. In particular, we will show that when either the central distortion constraint or the side distortion constraint is loose, the outer bound (12)-(13) is still tight. Let

$$k_l = (d_l^{-1} - \sigma_x^{-2})^{-1}, \quad l = 1, \dots, L \quad (45)$$

in *Lemma 5.1*. To simplify the derivation, we use $\underline{k}_w(\sigma_1^2)$ and $\sigma_j^2(\sigma_1^2)$, $j = 2, \dots, J$ to denote the functions of \underline{k}_w and σ_j^2 , $j = 2, \dots, J$ over σ_1^2 . From (43), we denote the upper bound of $\underline{k}_w(\sigma_1^2)$ as $\underline{k}_w^{up} = \left(\sum_{i=1}^L k_i^{-1} \right)^{-1}$. Considering the central distortion constraint, we introduce

$$\underline{k}_w^\diamond = (d_0^{-1} - \sigma_x^{-2})^{-1}.$$

We now consider two scenarios depending on the relationship between \underline{k}_w^{up} and \underline{k}_w^\diamond :

Scenario 1 ($\underline{k}_w^\diamond \geq \underline{k}_w^{up}$): The analysis for this scenario is trivial. It corresponds to the case that the central distortion constraint is loose. By letting

$$\sigma_j^2 = 0, \quad j = 1, \dots, J,$$

the optimal Gaussian test channel is thus specified. The resulting central distortion is

$$d_0^* = \sum_{i=1}^L d_i^{-1} - (L-1)\sigma_x^{-2}. \quad (46)$$

Further, there is $d_0^* \leq d_0$. In this situation, the optimal rate for each channel is actually the minimum single description rate, i.e. $R_l = \frac{1}{2} \log \frac{\sigma_x^2}{d_l}$, $k = 1, \dots, L$. From (14), it is straightforward that the Gaussian description scheme achieves the optimal weighted sum rate.

Scenario 2 ($\underline{k}_w^\diamond < \underline{k}_w^{up}$): From *Lemma 5.1*, it is ensured that there exists $\tilde{\sigma}_1^2 > 0$ such that the corresponding parameter $\underline{k}_w(\tilde{\sigma}_1^2)$ satisfies

$$\underline{k}_w(\tilde{\sigma}_1^2) = \underline{k}_w^\diamond.$$

In this situation, there are two possible outcomes when comparing $\sigma_J^2(\tilde{\sigma}_1^2)$ with σ_x^2 .

We first consider the case that $\sigma_J^2(\tilde{\sigma}_1^2) < \sigma_x^2$. From *Lemma 5.1*, it follows that

$$0 < \alpha_1 \tilde{\sigma}_1^2 < \alpha_2 \sigma_2^2(\tilde{\sigma}_1^2) < \dots < \alpha_J^2(\tilde{\sigma}_1^2) < \alpha_J \sigma_x^2.$$

We construct the Gaussian test channel using $(\tilde{\sigma}_1^2, \dots, \sigma_J^2(\tilde{\sigma}_1^2))$ and (45). The corresponding multiple description scheme using Gaussian descriptions satisfies the conditions of *Theorem 2.3*.

Thus, we obtain the optimal weighted sum rate.

Next we consider the case that $\sigma_J^2(\tilde{\sigma}_1^2) \geq \sigma_x^2$. This particular case actually corresponds to the situation that the individual distortion constraints associated with the weighting factor α_J are loose. To show that the outer bound (12) is still tight, the basic idea is to first construct a new d'_L smaller than or equal to d_L in (45), and then consider the new multiple description problem with distortion constraints $(d_1, \dots, d_{L-1}, d'_L, d_0)$. The key observation is that the resulting optimal sum rate (12) of the new multiple description problem is only a function of $(d_0, \dots, d_{L-1}, d_0)$, and is unrelated with d'_L or d_L . From *Lemma 5.1*, there exists $0 < \bar{\sigma}_1^2 \leq \tilde{\sigma}_1^2$ such that

$$\begin{cases} \sigma_J^2(\bar{\sigma}_1^2) = \sigma_x^2 \\ \underline{k}_w^\diamond \leq \underline{k}_w(\bar{\sigma}_1^2) \end{cases}.$$

In this situation, we have

$$\left(\underline{k}_w^\diamond + \sigma_J^2(\bar{\sigma}_1^2)\right)^{-1} = \left(\underline{k}_{\{1, \dots, M_1^{J-1}\}} + \sigma_J^2(\bar{\sigma}_1^2)\right)^{-1} + \sum_{i=1}^{m_J} \left(k_{M_1^{J-1}+i} + \sigma_J^2(\bar{\sigma}_1^2)\right)^{-1} + \lambda, \quad (47)$$

where $\lambda \geq 0$, and $\underline{k}_{\{1, \dots, M_1^{J-1}\}}$ is a function of $\bar{\sigma}_1^2$ as defined in *Lemma 5.1*. We use the *enhancement* technique [13] to find a k'_L such that $k'_L \leq k_L$ (correspondingly $d'_L \leq d_L$). Letting

$$(k_L + \sigma_J^2(\bar{\sigma}_1^2))^{-1} + \lambda = (k'_L + \sigma_J^2(\bar{\sigma}_1^2))^{-1}$$

and using the fact that $\sigma_x^2 = \sigma_J^2(\bar{\sigma}_1^2)$, we arrive at

$$k'_L = [(k_L + \sigma_x^2)^{-1} + \lambda]^{-1} - \sigma_x^2. \quad (48)$$

The verification of $k'_L \leq k_L$ is straightforward. Equ. (47) can be rewritten in terms of k'_L as

$$\left(\underline{k}_0^\diamond + \sigma_J^2(\bar{\sigma}_1^2)\right)^{-1} = \left(\underline{k}_{\{1, \dots, M_1^{J-1}\}} + \sigma_J^2(\bar{\sigma}_1^2)\right)^{-1} + \sum_{i=1}^{m_J-1} \left(k_{M_1^{J-1}+i} + \sigma_J^2(\bar{\sigma}_1^2)\right)^{-1} + (k'_L + \sigma_J^2(\bar{\sigma}_1^2))^{-1}. \quad (49)$$

From Lemma 2.5 and Lemma 5.1, we conclude that the Gaussian test channel built from $(k_1, \dots, k_{L-1}, k'_L)$ and $(\bar{\sigma}_1^2, \dots, \sigma_J^2(\bar{\sigma}_1^2))$ is valid. The resulting distortions are $(d_1, \dots, d_{L-1}, d'_L, d_0)$, where $0 < d'_L \leq d_L$.

We now show that the optimal weighted sum rate (12) of the above Gaussian description scheme is unrelated with d'_L . The idea is to construct a sequence $d_0(\epsilon)$ such that $\sigma_J^2(\bar{\sigma}_1^2(\epsilon)) < \sigma_x^2$, thus allowing the application of Theorem 2.3. Then by letting $\epsilon \rightarrow 0$, the argument is straightforward. Let

$$d_0(\epsilon) = ((\underline{k}_w^\diamond + \epsilon)^{-1} + \sigma_x^2)^{-1}, \quad (50)$$

where ϵ is chosen such that $\epsilon > 0$ and $(\underline{k}_w^\diamond + \epsilon)^{-1} > \sum_{i=1}^{L-1} k_i^{-1} + k'_L{}^{-1}$. Taking $(k_1, \dots, k_{L-1}, k'_L)$ in Lemma 5.1, we obtain

$$\begin{aligned} \underline{k}_w^\diamond + \epsilon &= \underline{k}_w(\bar{\sigma}_1^2(\epsilon)), \\ \sigma_J^2(\bar{\sigma}_1^2(\epsilon)) &< \sigma_x^2, \quad \epsilon \neq 0. \end{aligned}$$

Considering the multiple description problem with distortion constraints $(d_1, \dots, d'_L, d_0(\epsilon))$, the situation corresponds to the one discussed in the first case. As long as $\epsilon \neq 0$, the optimal weighted sum rate of the modified distortion multiple description problem is known, which is given by (12). Particularly, the expressions of the optimal $n_j(\epsilon)$ (n_j is the scalar version of N_j in (12)), $j = 1, \dots, J$ take the form (see Proposition 2.6)

$$n_j(\epsilon) = \frac{\sigma_x^2}{\sigma_x^2 / \sigma_j^2(\bar{\sigma}_1^2(\epsilon)) - 1}, \quad j = 1, \dots, J. \quad (51)$$

Taking the limit of ϵ , i.e. $\epsilon \downarrow 0$, there are $d_0(\epsilon) \downarrow d_0$, $\sigma_J^2(\bar{\sigma}_1^2(\epsilon)) \uparrow \sigma_x^2$ and $n_J(\epsilon) \uparrow +\infty$. Due to the limiting operation $n_J(\epsilon) \uparrow +\infty$, the side distortion d'_L does not contribute to the resulting optimal weighted sum rate (12). In other words, the optimal weighted sum rate (12) is only a function of $(d_1, \dots, k_{L-1}, d_0)$. Thus, we show that the outer bound (12) is still tight for the original multiple description problem with distortion constraints (d_1, \dots, d_L, d_0) (i.e., the individual side distortions are loose).

In summary, for a scalar Gaussian source, the Gaussian description scheme fully characterizes the rate region. When the central distortion constraint is trivial ($\underline{k}_w^\diamond \geq \underline{k}_w^{up}$), the shape of the rate region belongs to the class of contra-polymatroids [14]. If it is not the case, one can always

construct an optimal Gaussian test channel satisfying the distortion constraint. This is to say that when $\underline{k}_w^\diamond < \underline{k}_w^{up}$, there always exists $0 < \alpha_1 \sigma_1^2 < \dots, < \alpha_J \sigma_J^2 \leq \alpha_J \sigma_x^2$ such that

$$\alpha_1 \sum_{i=1}^{m_1} (\sigma_1^2 + k_i)^{-1} = \frac{\alpha_2}{\sigma_1^2} - \frac{\alpha_1 - \alpha_2}{\sigma_2^2 - \sigma_1^2}, \quad (52)$$

$$\alpha_j \sum_{i=1}^{m_j} (\sigma_j^2 + k_{M_1^{j-1}+i})^{-1} + \frac{\alpha_{j-1}}{\sigma_j^2} - \frac{\alpha_{j-1} - \alpha_j}{\sigma_j^2 - \sigma_{j-1}^2} = \frac{\alpha_{j+1}}{\sigma_j^2} - \frac{\alpha_j - \alpha_{j+1}}{\sigma_{j+1}^2 - \sigma_j^2}, \quad j = 2, \dots, J-1, \quad (53)$$

$$\lambda + \alpha_J \sum_{i=1}^{m_J} (\sigma_J^2 + k_{M_1^{J-1}+i})^{-1} + \frac{\alpha_{J-1}}{\sigma_J^2} - \frac{\alpha_{J-1} - \alpha_J}{\sigma_J^2 - \sigma_{J-1}^2} = \alpha_J (\sigma_J^2 + k_0^\diamond)^{-1}, \quad (54)$$

where

$$\lambda(\sigma_x^2 - \sigma_J^2) = 0, \quad \lambda \geq 0, \quad (55)$$

and $k_l, l = 1, \dots, L$ are as defined in (45). The conditions (52)-(55) follow from (39)-(40), (45) and (47). Further, the solution to (52)-(55) is unique.

The solution to (52)-(55) can be interpreted as a solution to an optimization problem. We first introduce parameters $y_j > 0, j = 1, \dots, J$ where $\sigma_j^2 = \sum_{i=1}^j y_i$. Define a new function $F(y_1, \dots, y_J)$ as

$$\begin{aligned} F(y_1, \dots, y_J) = & \alpha_J \log \left(\underline{k}_w^\diamond + \sum_{d=1}^J y_d \right) - \sum_{j=1}^J \alpha_j \sum_{i=1}^{m_j} \log \left(\sum_{d=1}^j y_d + k_{M_1^{j-1}+i} \right) + \sum_{j=1}^{J-1} (\alpha_j - \alpha_{j+1}) \log y_{j+1} \\ & - \sum_{j=2}^{J-1} (\alpha_{j-1} - \alpha_{j+1}) \log \left(\sum_{i=1}^j y_i \right) + \alpha_2 \log y_1 - \alpha_{J-1} \log \left(\sum_{i=1}^J y_i \right). \end{aligned} \quad (56)$$

Consider the following optimization problem:

$$\max_{\{y_j\}_{j=1}^J} F(y_1, \dots, y_J) \quad \text{subject to} \quad \begin{cases} 0 < y_j, & j = 1, \dots, J \\ \sum_{j=1}^J y_j \leq \sigma_x^2 \end{cases}. \quad (57)$$

Using the concept of *Lagrange duality* [15, Chapter 5] to handle the constraint $\sum_{j=1}^J y_j \leq \sigma_x^2$, the corresponding Karush-Kuhn-Tucker (KKT) conditions take the form

$$\frac{\partial F}{\partial y_j} - \gamma = 0, \quad j = 1, \dots, J \quad (58)$$

$$\gamma(\sigma_x^2 - \sum_{j=1}^J y_j) = 0, \quad (59)$$

$$\sum_{j=1}^J y_j \leq \sigma_x^2 \text{ and } y_j > 0, j = 1, \dots, J. \quad (60)$$

One can show that (58)-(60) are equivalent to (52)-(55) by plugging $y_j = \sigma_j^2 - \sigma_{j-1}^2$ and $\gamma = \lambda$ into the expressions. Thus, instead of solving a set of equations, we can equivalently consider an optimization problem. The sufficient condition making the KKT conditions hold is $\underline{k}_w^\diamond < \underline{k}_w^{up}$ as discussed in **Scenario 2**.

VI. CONCLUSION

We have addressed the rate region of the multiple description problem with respect to individual and central distortion constraints for a vector Gaussian source. An outer bound was derived for the considered rate region. In the special case of a scalar Gaussian source, the lower bound was shown to be tight.

The work in [8] treated the special case that all the weighting factors in (3) are identical (i.e., $J = 1$), which corresponds to minimizing the sum rate. In this particular case, it was previously shown that the outer bound given by (13), remains tight even if some of the distortion constraints are loose. Thus, the optimal sum rate for individual and central receivers has been completely characterized [8]. While we have not been able to prove similar result for the weighted sum rate case (where $J > 1$), we believe that the outer bound given by (12), remains tight even if some of the distortion constraints are loose.

APPENDIX A

USEFUL MATRIX LEMMAS

Lemma A.1 (Matrix Inversion Lemma): [16, Theorem 2.5] Let \mathbf{A} be an $m \times m$ nonsingular matrix and \mathbf{B} be an $n \times n$ nonsingular matrix and let \mathbf{C} and \mathbf{D} be $m \times n$ and $n \times n$ matrices, respectively. If the matrix $\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{D}$ is nonsingular, then

$$(\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1} + \mathbf{D}\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{D}\mathbf{A}^{-1}. \quad (61)$$

Lemma A.2: [16, Theorems 6.8 and 6.9] Let \mathbf{A} and \mathbf{B} be positive definite matrices such that $\mathbf{A} \succ \mathbf{B}$ ($\mathbf{A} \succeq \mathbf{B}$). Then

$$|\mathbf{A}| \succ |\mathbf{B}| \quad (|\mathbf{A}| \succeq |\mathbf{B}|)$$

$$\mathbf{A}^{-1} \prec \mathbf{B}^{-1} \quad (\mathbf{A}^{-1} \preceq \mathbf{B}^{-1}). \quad (62)$$

APPENDIX B

PROOF OF LEMMA 3.1

The main tool used in the proof is the generalized Costa's entropy-power inequality (EPI) developed by R. Liu et al. in [17]. Specifically, we use the conditional version of the generalized Costa's EPI, which was shown to be a simple extension of the generalized Costa's EPI [17]. To make the work complete, we present the result in a lemma below.

Lemma B.1 (Theorem 1, Corollary 1 in [17]): Let \mathbf{z} be a Gaussian random n -vector with a positive definite covariance matrix \mathbf{C} , and let \mathbf{A} be an $n \times n$ real symmetric matrix such that $\mathbf{0} \preceq \mathbf{A} \preceq \mathbf{I}_n$. Then

$$\exp \left[\frac{2}{n} h(\mathbf{p} + \mathbf{A}^{\frac{1}{2}} \mathbf{z} | \mathbf{v}) \right] \geq |\mathbf{I} - \mathbf{A}|^{\frac{1}{n}} \exp \left[\frac{2}{n} h(\mathbf{p} | \mathbf{v}) \right] + |\mathbf{A}|^{\frac{1}{n}} \exp \left[\frac{2}{n} h(\mathbf{p} + \mathbf{z} | \mathbf{v}) \right] \quad (63)$$

for any \mathbf{v} and n -vector \mathbf{p} independent of \mathbf{z} . The equality holds if (\mathbf{p}, \mathbf{v}) are jointly Gaussian with a conditional covariance matrix $\mathbf{B} = \text{Cov}[\mathbf{p} | \mathbf{v}]$ such that $\mathbf{B} - \mathbf{A}\mathbf{B}$ and $\mathbf{B} + \mathbf{A}^{\frac{1}{2}}\mathbf{C}\mathbf{A}^{\frac{1}{2}}$ are proportional.

Next we are in a position to prove the lemma. The basic idea is as follows. we first apply a linear transformation to the nN -dimensional random vectors \mathbf{x}^n , \mathbf{z}_1^n and \mathbf{z}_2^n such that after transformation the random vectors corresponding to \mathbf{z}_1^n and \mathbf{z}_2^n have diagonal covariance matrices. The Costa's EPI (63) is then used to prove an extremal entropy inequality in the transform domain. Finally, the upper-bound inequality (19) is obtained by converting the extremal entropy inequality to the original domain.

We consider diagonalizing \mathbf{N}_1 and \mathbf{N}_2 simultaneously. Using the fact that \mathbf{N}_1 and \mathbf{N}_2 are positive definite, it is known [18] that there exists an invertible matrix \mathbf{U} such that

$$\mathbf{U}^t \mathbf{N}_1 \mathbf{U} = \mathbf{\Lambda}_1 \text{ and } \mathbf{U}^t \mathbf{N}_2 \mathbf{U} = \mathbf{\Lambda}_2, \quad (64)$$

where Λ_1 and Λ_2 are positive definite diagonal matrices. From the assumption on N_1 and N_2 , it is immediate that $\Lambda_1 \prec \Lambda_2$. Let

$$\mathbf{y}_1^n = (\mathbf{I}_n \otimes \mathbf{U}^t)(\mathbf{x}^n + \mathbf{z}_1^n) \quad (65)$$

$$\text{and } \mathbf{y}_2^n = (\mathbf{I}_n \otimes \mathbf{U}^t)(\mathbf{x}^n + \mathbf{z}_2^n). \quad (66)$$

Note that by applying the linear transformation in (65)-(66), the resulting random vectors corresponding to \mathbf{z}_1^n and \mathbf{z}_2^n have independent components. To simplify the notation in the derivation, we introduce

$$\tilde{\mathbf{x}}^n = (\mathbf{I}_n \otimes \mathbf{U}^t)\mathbf{x}^n \quad (67)$$

$$\text{and } \tilde{\mathbf{z}}^n = (\mathbf{I}_n \otimes \mathbf{U}^t)\mathbf{z}_2^n. \quad (68)$$

Consequently we have $\text{Cov}[\tilde{\mathbf{z}}^n] = \mathbf{I}_n \otimes \Lambda_2$. The two variables \mathbf{y}_1^n and \mathbf{y}_2^n can be equivalently written as

$$\mathbf{y}_2^n = \tilde{\mathbf{x}}^n + \tilde{\mathbf{z}}^n \quad (69)$$

$$\mathbf{y}_1^n = \tilde{\mathbf{x}}^n + (\mathbf{I}_n \otimes \mathbf{A}^{\frac{1}{2}})\tilde{\mathbf{z}}^n, \quad (70)$$

where $\mathbf{A} = \Lambda_1 \Lambda_2^{-1}$ satisfies $\mathbf{0} \prec \mathbf{A} \prec \mathbf{I}$. Considering the quantity $h(\mathbf{y}_1^n | \mathbf{v})$, it is immediate from Lemma B.1 that

$$h(\mathbf{y}_1^n | \mathbf{v}) \geq \frac{Nn}{2} \log \left[|\mathbf{I}_N - \mathbf{A}|^{\frac{1}{N}} \exp \left[\frac{2}{Nn} h(\tilde{\mathbf{x}}^n | \mathbf{v}) \right] + |\mathbf{A}|^{\frac{1}{N}} \exp \left[\frac{2}{Nn} h(\mathbf{y}_2^n | \mathbf{v}) \right] \right]. \quad (71)$$

We now prove an extremal entropy inequality in the transform domain. By using (71), we have

$$\begin{aligned} & \mu_2 h(\mathbf{y}_2^n | \mathbf{v}) - \mu_1 h(\mathbf{y}_1^n | \mathbf{v}) + (\mu_1 - \mu_2) h(\tilde{\mathbf{x}}^n | \mathbf{v}) \\ & \leq \mu_2 h(\mathbf{y}_2^n | \mathbf{v}) - \frac{\mu_1 Nn}{2} \log \left[|\mathbf{I}_N - \mathbf{A}|^{\frac{1}{N}} \exp \left[\frac{2}{Nn} h(\tilde{\mathbf{x}}^n | \mathbf{v}) \right] + |\mathbf{A}|^{\frac{1}{N}} \exp \left[\frac{2}{Nn} h(\mathbf{y}_2^n | \mathbf{v}) \right] \right] + (\mu_1 - \mu_2) h(\tilde{\mathbf{x}}^n | \mathbf{v}) \\ & = \mu_2 I(\tilde{\mathbf{z}}^n; \tilde{\mathbf{x}}^n + \tilde{\mathbf{z}}^n | \mathbf{v}) - \frac{\mu_1 Nn}{2} \log \left[|\mathbf{I}_N - \mathbf{A}|^{\frac{1}{N}} + |\mathbf{A}|^{\frac{1}{N}} \exp \left[\frac{2}{Nn} I(\tilde{\mathbf{z}}^n; \tilde{\mathbf{x}}^n + \tilde{\mathbf{z}}^n | \mathbf{v}) \right] \right]. \end{aligned} \quad (72)$$

From (72), the function

$$f(t) = \mu_2 t - \frac{\mu_1 N n}{2} \log \left(|\mathbf{I}_N - \mathbf{A}|^{\frac{1}{N}} + |\mathbf{A}|^{\frac{1}{N}} \exp \left[\frac{2}{N n} t \right] \right) \quad (73)$$

is concave in t and has a global maxima at

$$t^* = \frac{N n}{2} \log \left(\frac{\mu_2 |\mathbf{A}^{-1} - \mathbf{I}_N|^{\frac{1}{N}}}{\mu_1 - \mu_2} \right). \quad (74)$$

Combining (72) and (74) produces

$$\mu_2 h(\mathbf{y}_2^n | \mathbf{v}) - \mu_1 h(\mathbf{y}_1^n | \mathbf{v}) + (\mu_1 - \mu_2) h(\tilde{\mathbf{x}}^n | \mathbf{v}) \leq \frac{\mu_2 n}{2} \log \frac{\mu_2^N |\mathbf{\Lambda}_2 - \mathbf{\Lambda}_1|}{(\mu_1 - \mu_2)^N |\mathbf{\Lambda}_1|} - \frac{\mu_1 n}{2} \log \frac{\mu_1^N |\mathbf{\Lambda}_2 - \mathbf{\Lambda}_1|}{(\mu_1 - \mu_2)^N |\mathbf{\Lambda}_2|}. \quad (75)$$

The equality in (75) holds if $(\tilde{\mathbf{x}}^n, \mathbf{v})$ satisfies the equality condition in (71) imposed by the generalized Costa's EPI and the optimality condition (74). From Lemma B.1, the equality condition to (71) can be mathematically written as

$$(\mathbf{I}_{nN} - \mathbf{I}_n \otimes \mathbf{A}) \text{Cov}(\tilde{\mathbf{x}}^n | \mathbf{v}) = c \left[\text{Cov}(\tilde{\mathbf{x}}^n | \mathbf{v}) + \mathbf{I}_n \otimes (\mathbf{A}^{1/2} \mathbf{\Lambda}_2 \mathbf{A}^{1/2}) \right], \quad (76)$$

where c is a scalar variable. By combining (74) and (76), it is found that the above two conditions are equivalent to the fact that $(\tilde{\mathbf{x}}^n, \mathbf{v})$ are jointly Gaussian with a conditional covariance matrix

$$\text{Cov}[\tilde{\mathbf{x}}^n | \mathbf{v}] = \mathbf{I}_n \otimes [(\mu_1 - \mu_2) \mathbf{\Lambda}_2 (\mu_2 \mathbf{\Lambda}_2 - \mu_1 \mathbf{\Lambda}_1)^{-1} \mathbf{\Lambda}_1]. \quad (77)$$

The final step is to convert the inequality (75) to an inequality in the original domain. In other words, we relate (75)-(77) to (19)-(20). From (65)-(67), we arrive at

$$h(\mathbf{y}_i^n | \mathbf{v}) = h(\mathbf{x}^n + \mathbf{z}_i^n | \mathbf{v}) + n \log |\mathbf{U}^t|, \quad i = 1, 2 \quad (78)$$

$$h(\tilde{\mathbf{x}}^n | \mathbf{v}) = h(\mathbf{x}^n | \mathbf{v}) + n \log |\mathbf{U}^t| \quad (79)$$

$$\text{Cov}[\tilde{\mathbf{x}}^n | \mathbf{v}] = [\mathbf{I}_n \otimes \mathbf{U}^t] \text{Cov}[\mathbf{x}^n | \mathbf{v}] [\mathbf{I}_n \otimes \mathbf{U}]. \quad (80)$$

By plugging (78)-(80) into (75)-(77), the derivation of (19)-(20) is then straightforward. The condition $\mathbf{0} \prec \mu_1 \mathbf{N}_1 \prec \mu_2 \mathbf{N}_2$ imposed on \mathbf{N}_1 and \mathbf{N}_2 is to ensure that the conditional covariance matrix $\text{Cov}[\mathbf{x}^n | \mathbf{v}]$ in (20) is positive definite.

APPENDIX C

PROOF OF LEMMA 2.4

The proof is essentially the same for every $j = 1, \dots, J$. Thus we only consider the derivation for a particular j . we first let $\mathbf{A}_j \succ \mathbf{0}$. By using Lemma A.1, there is

$$\begin{aligned}
& \left[\mathbf{A}_j^{-1} + (\mathbf{I}_N, \dots, \mathbf{I}_N) \mathbf{K}_{\{1, \dots, M_1^j\}}^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \right]^{-1} \\
&= \mathbf{A}_j - \mathbf{A}_j (\mathbf{I}_N, \dots, \mathbf{I}_N) \left[\mathbf{K}_{\{1, \dots, M_1^j\}} + (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \mathbf{A}_j (\mathbf{I}_N, \dots, \mathbf{I}_N) \right]^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \mathbf{A}_j \\
&= \mathbf{A}_j - \mathbf{A}_j \left[(\mathbf{I}_N, \dots, \mathbf{I}_N) \left(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right)^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \right. \\
&\quad \left. + \sum_{i=1}^{m_j} \left(\mathbf{K}_{M_1^{j-1}+i} + \mathbf{A}_j \right)^{-1} \right] \mathbf{A}_j. \tag{81}
\end{aligned}$$

By rearranging the items in (81), we have

$$\begin{aligned}
& (\mathbf{I}_N, \dots, \mathbf{I}_N) \mathbf{K}_{\{1, \dots, M_1^j\}}^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \\
&= \left[\mathbf{A}_j - \mathbf{A}_j \left((\mathbf{I}_N, \dots, \mathbf{I}_N) \left(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right)^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t + \right. \right. \\
&\quad \left. \left. \sum_{i=1}^{m_j} \left(\mathbf{K}_{M_1^{j-1}+i} + \mathbf{A}_j \right)^{-1} \right) \mathbf{A}_j \right]^{-1} - \mathbf{A}_j^{-1} \\
&= \left[-\mathbf{A}_j + \left((\mathbf{I}_N, \dots, \mathbf{I}_N) \left(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right)^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t + \right. \right. \\
&\quad \left. \left. \sum_{i=1}^{m_j} \left(\mathbf{K}_{M_1^{j-1}+i} + \mathbf{A}_j \right)^{-1} \right) \right]^{-1}. \tag{82}
\end{aligned}$$

With the definition of $\underline{\mathbf{K}}_S$ in (8), (82) can be rewritten as

$$\left(\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}} + \mathbf{A}_j \right)^{-1} = (\mathbf{I}_N, \dots, \mathbf{I}_N) \left(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right)^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t + \sum_{i=1}^{m_j} \left(\mathbf{K}_{M_1^{j-1}+i} + \mathbf{A}_j \right)^{-1}. \tag{83}$$

We now consider simplifying (83) further. By using Lemma A.1, there is

$$\begin{aligned}
& \left[(-\mathbf{A}_j)^{-1} + (\mathbf{I}_N, \dots, \mathbf{I}_N) \left(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right)^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \right]^{-1} \\
&= -\mathbf{A}_j - (-\mathbf{A}_j) (\mathbf{I}_N, \dots, \mathbf{I}_N) \mathbf{K}_{\{1, \dots, M_1^{j-1}\}}^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t (-\mathbf{A}_j). \tag{84}
\end{aligned}$$

Similarly to the derivation of (82) from (81), (84) can be rewritten as

$$\begin{aligned}
& (\mathbf{I}_N, \dots, \mathbf{I}_N) \left(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right)^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \\
&= \left[-\mathbf{A}_j - \mathbf{A}_j (\mathbf{I}_N, \dots, \mathbf{I}_N) \mathbf{K}_{\{1, \dots, M_1^{j-1}\}}^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \mathbf{A}_j \right]^{-1} + \mathbf{A}_j^{-1} \\
&= \left[\left((\mathbf{I}_N, \dots, \mathbf{I}_N) \mathbf{K}_{\{1, \dots, M_1^{j-1}\}}^{-1} (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \right)^{-1} + \mathbf{A}_j \right]^{-1} \\
&= \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{A}_j \right)^{-1}. \tag{85}
\end{aligned}$$

Combining (83) and (85) produces (17). For the case that $\mathbf{A}_j \succeq \mathbf{0}$, there exists $\delta > 0$ such that $\mathbf{A}_j + \epsilon \mathbf{I}_N \succ \mathbf{0}$ and the corresponding $\mathbf{K}_{\{1, \dots, M_1^j\}} \succ \mathbf{0}$ for $\epsilon \in (0, \delta)$, which supports the previous argument. By letting $\epsilon \rightarrow 0^+$, we obtain the result. The proof is complete.

APPENDIX D

PROOF OF LEMMA 5.1

We prove the lemma using induction argument. We first consider the monotonicity properties of σ_2^2 and $\underline{k}_{\{1, \dots, M_1^1\}}$ over σ_1^2 , referred to as **Case 1**. Later, we extend the analysis to the general case. The proof is rather long. We present the proof in several steps.

Case 1 ($j = 1$): In order to discuss the properties of σ_2^2 and $\underline{k}_{\{1, \dots, M_1^1\}}$ as functions of σ_1^2 , we first study the support region of σ_1^2 such that $\sigma_2^2 > \sigma_1^2$. By combining (39) and (40), we obtain

$$\frac{\alpha_2}{\alpha_1 \sigma_1^2} + \frac{\alpha_2 - \alpha_1}{\alpha_1 (\sigma_2^2 - \sigma_1^2)} = \sum_{i=1}^{m_1} (\sigma_1^2 + k_i)^{-1}.$$

In order that $\sigma_2^2 > \sigma_1^2$, there is

$$\frac{\alpha_2}{\alpha_1 \sigma_1^2} > \sum_{i=1}^{m_1} (\sigma_1^2 + k_i)^{-1} \stackrel{\sigma_1^2 > 0}{\Rightarrow} \sum_{i=1}^{m_1} \frac{k_i}{\sigma_1^2 + k_i} > m_1 - \frac{\alpha_2}{\alpha_1}.$$

Letting

$$f(\sigma_1^2) = \sum_{i=1}^{m_1} \frac{k_i}{\sigma_1^2 + k_i} - \left(m_1 - \frac{\alpha_2}{\alpha_1} \right), \tag{86}$$

we have $f(0) > 0$ and $f(+\infty) < 0$. By using intermediate value theorem [19, p. 48], it is

obvious that there exists $\hat{\sigma}_{\{1\}}^2 > 0$ such that

$$0 < \sigma_1^2 < \hat{\sigma}_{\{1\}}^2 \Rightarrow \sigma_2^2 > \sigma_1^2 \quad (87)$$

$$\text{and} \quad \begin{cases} \sigma_2^2 \rightarrow \sigma_1^2 \text{ as } \sigma_1^2 \rightarrow 0 \\ \sigma_2^2 \rightarrow +\infty \text{ as } \sigma_1^2 \rightarrow \hat{\sigma}_{\{1\}}^2 \end{cases}. \quad (88)$$

Further, by analyzing (39)-(40) under the situation (87)-(88), we find that $\underline{k}_{\{1, \dots, M_1^1\}}$ can be bounded as

$$A_1 < \underline{k}_{\{1, \dots, M_1^1\}} < \left(\sum_{i=1}^{m_1} k_i^{-1} \right)^{-1}, \quad (89)$$

where $A_1 > 0$, and is determined by $\hat{\sigma}_{\{1\}}^2$. The key observation here is that $\underline{k}_{\{1, \dots, M_1^1\}}$ is bounded away from 0 and $+\infty$. The particular value A_1 is not important. Using the fact that $\underline{k}_{\{1, \dots, M_1^1\}} > 0$ and (40), it is straightforward that $\alpha_1 \sigma_1^2 < \alpha_2 \sigma_2^2$.

We now discuss the monotonicity properties of $\underline{k}_{\{1, \dots, M_1^1\}}$ and σ_2^2 over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1\}}^2)$. Calculating the differentiation of (39) and (40) with respect to σ_1^2 produces

$$\left(\frac{d\underline{k}_{\{1, \dots, M_1^1\}}}{d\sigma_1^2} + 1 \right) (\underline{k}_{\{1, \dots, M_1^1\}} + \sigma_1^2)^{-2} = \sum_{i=1}^{m_1} (\sigma_1^2 + k_i)^{-2} \quad (90)$$

$$\text{and} \quad -\frac{d\underline{k}_{\{1, \dots, M_1^1\}}}{d\sigma_1^2} k_{\{1, \dots, M_1^1\}}^{-2} = -\frac{\alpha_2}{(\alpha_1 - \alpha_2)\sigma_1^4} + \frac{\alpha_1}{(\alpha_1 - \alpha_2)\sigma_2^4} \frac{d\sigma_2^2}{d\sigma_1^2}. \quad (91)$$

Combining (39) and (90) yields

$$\frac{d\underline{k}_{\{1, \dots, M_1^1\}}}{d\sigma_1^2} (\underline{k}_{\{1, \dots, M_1^1\}} + \sigma_1^2)^{-2} = \sum_{i=1}^{m_1} (\sigma_1^2 + k_i)^{-2} - \left(\sum_{i=1}^{m_1} (\sigma_1^2 + k_i)^{-1} \right)^2. \quad (92)$$

It is easily seen from (91)-(92) that $\frac{d\underline{k}_{\{1, \dots, M_1^1\}}}{d\sigma_1^2} < 0$ and $\frac{d\sigma_2^2}{d\sigma_1^2} > 0$ when $\sigma_1^2 \in (0, \hat{\sigma}_{\{1\}}^2)$. Thus, $\underline{k}_{\{1, \dots, M_1^1\}}$ is monotonically decreasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1\}}^2)$, and bounded away from 0 and $+\infty$. The parameter σ_2^2 is monotonically increasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1\}}^2)$, and can take any positive real value. The parameters σ_1^2 and σ_2^2 satisfies the inequality $\alpha_1 \sigma_1^2 < \alpha_2 \sigma_2^2$ for any $\sigma_1^2 \in (0, \hat{\sigma}_{\{1\}}^2)$.

Case 2 ($j = 2, \dots, J-1$): We now study the monotonicity properties of σ_{j+1}^2 and $k_{\{1, \dots, M_1^j\}}$ over σ_1^2 by assuming that some prior information about σ_j^2 and $k_{\{1, \dots, M_1^{j-1}\}}$ is known. Specifically, we assume that σ_j^2 is monotonically increasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1, \dots, j-1\}}^2)$, and can take any positive real value. The parameter $\hat{\sigma}_{\{1, \dots, j-1\}}^2$ is determined using the constraint that $\sigma_j^2 > \sigma_{j-1}^2$. Also we

assume $\underline{k}_{\{1, \dots, M_1^{j-1}\}}$ is monotonically decreasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1, \dots, j-1\}}^2)$, and is bounded by $A_{j-1} < \underline{k}_{\{1, \dots, M_1^{j-1}\}} < \left(\sum_{i=1}^{M_1^{j-1}} k_i^{-1}\right)^{-1}$.

Similarly to the analysis for the case of $j = 1$, we first study the support region of σ_1^2 such that $\sigma_{j+1}^2 > \sigma_j^2$. Again by combining (39)-(40), there is

$$\frac{\alpha_{j+1}}{\alpha_j \sigma_j^2} + \frac{\alpha_{j+1} - \alpha_j}{\alpha_j (\sigma_{j+1}^2 - \sigma_j^2)} = \sum_{j=1}^{m_j} \left(\sigma_j^2 + k_{M_1^{j-1}+i} \right)^{-1} + \left(\sigma_j^2 + k_{M_1^{j-1}} \right)^{-1}. \quad (93)$$

We consider the inequality

$$\begin{aligned} \frac{\alpha_{j+1}}{\alpha_j \sigma_j^2} &> \sum_{i=1}^{m_j} \left(\sigma_j^2 + k_{M_1^{j-1}+i} \right)^{-1} + \left(\sigma_j^2 + \underline{k}_{\{1, \dots, M_1^{j-1}\}} \right)^{-1} \\ \stackrel{\sigma_j^2 > 0}{\Rightarrow} \sum_{i=1}^{m_j} \frac{k_{M_1^{j-1}+i}}{\sigma_j^2 + k_{M_1^{j-1}+i}} + \frac{\underline{k}_{\{1, \dots, M_1^{j-1}\}}}{\sigma_j^2 + \underline{k}_{\{1, \dots, M_1^{j-1}\}}} &> m_j + 1 - \frac{\alpha_{j+1}}{\alpha_j}. \end{aligned}$$

Note that σ_j^2 and $\underline{k}_{\{1, \dots, M_1^{j-1}\}}$ are varying simultaneously along with σ_1^2 . Thus a direct extension of the analysis for **Case 1** is unapplicable here. Fortunately, it is known from the assumption that $\underline{k}_{\{1, \dots, M_1^{j-1}\}}$ is bounded away from 0 and $+\infty$ when $\sigma_1^2 \in (0, \hat{\sigma}_{\{1, \dots, j-1\}}^2)$. By using the above observation and the intermediate value theorem as in **Case 1**, we conclude that there exists $0 < \hat{\sigma}_{\{1, \dots, j\}}^2 < \hat{\sigma}_{\{1, \dots, j-1\}}^2$ such that

$$0 < \sigma_1^2 < \hat{\sigma}_{\{1, \dots, j\}}^2 \Rightarrow \sigma_{j+1}^2 > \sigma_j^2 \quad (94)$$

$$\text{and} \quad \begin{cases} \sigma_{j+1}^2 \rightarrow \sigma_j^2 \text{ as } \sigma_1^2 \rightarrow 0 \\ \sigma_{j+1}^2 \rightarrow +\infty \text{ as } \sigma_1^2 \rightarrow \hat{\sigma}_{\{1, \dots, j\}}^2 \end{cases} \quad (95)$$

The analysis implies that when the constraint $\sigma_{j+1}^2 > \sigma_j^2$ is imposed, the support region of σ_1^2 such that $\sigma_j^2 > \sigma_{j-1}^2$ becomes narrow, but still exists. Again by studying (39)-(40) under the situation (94) and (95), we arrive at

$$A_j < \underline{k}_{\{1, \dots, M_1^j\}} < \left(\sum_{i=1}^{M_1^j} k_i^{-1} \right)^{-1}, \quad (96)$$

where $A_j > 0$, and is determined by $\hat{\sigma}_{\{1, \dots, j\}}^2$. Using the fact that $\underline{k}_{\{1, \dots, M_1^j\}} > 0$ and (40), we obtain $\alpha_j \sigma_j^2 < \alpha_{j+1} \sigma_{j+1}^2$ for any $\sigma_1^2 \in (0, \hat{\sigma}_{\{1, \dots, j\}}^2)$.

Next we consider the monotonicity of σ_{j+1}^2 and $\underline{k}_{\{1, \dots, M_1^{j+1}\}}$ over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1, \dots, j\}}^2)$. Calculating

the differentiation of (39)-(40) w.r.t. σ_1^2 produces

$$\begin{aligned} \left(\frac{dk_{\{1,\dots,M_1^j\}}}{d\sigma_1^2} + \frac{d\sigma_j^2}{d\sigma_1^2} \right) \left(\underline{k}_{\{1,\dots,M_1^j\}} + \sigma_j^2 \right)^{-2} &= \sum_{i=1}^{m_j} \frac{d\sigma_j^2}{d\sigma_1^2} \left(\sigma_j^2 + k_{M_1^{j-1}+i} \right)^{-2} \\ &\quad + \left(\frac{d\sigma_j^2}{d\sigma_1^2} + \frac{dk_{\{1,\dots,M_1^{j-1}\}}}{d\sigma_1^2} \right) \left(\sigma_j^2 + \underline{k}_{\{1,\dots,M_1^{j-1}\}} \right)^{-2} \end{aligned} \quad (97)$$

$$\text{and } -\frac{dk_{\{1,\dots,M_1^j\}}}{d\sigma_1^2} \underline{k}_{\{1,\dots,M_1^j\}}^{-2} = -\frac{d\sigma_j^2}{d\sigma_1^2} \frac{\alpha_{j+1}}{(\alpha_j - \alpha_{j+1})\sigma_j^4} + \frac{d\sigma_{j+1}^2}{d\sigma_1^2} \frac{\alpha_j}{(\alpha_j - \alpha_{j+1})\sigma_{j+1}^4}. \quad (98)$$

Similarly to that of **Case 1**, By combining (39) and (97), we have

$$\begin{aligned} \frac{dk_{\{1,\dots,M_1^j\}}}{d\sigma_1^2} \left(\underline{k}_{\{1,\dots,M_1^j\}} + \sigma_j^2 \right)^{-2} &= \frac{dk_{\{1,\dots,M_1^{j-1}\}}}{d\sigma_1^2} \left(\sigma_j^2 + \underline{k}_{\{1,\dots,M_1^{j-1}\}} \right)^{-2} + \frac{d\sigma_j^2}{d\sigma_1^2} \left[\sum_{i=1}^{m_j} \left(\sigma_j^2 + k_{M_1^{j-1}+i} \right)^{-2} \right. \\ &\quad \left. + \left(\sigma_j^2 + \underline{k}_{\{1,\dots,M_1^{j-1}\}} \right)^{-2} - \left(\sum_{i=1}^{m_j} \left(\sigma_j^2 + k_{M_1^{j-1}+i} \right)^{-1} + \left(\sigma_j^2 + \underline{k}_{\{1,\dots,M_1^{j-1}\}} \right)^{-1} \right)^2 \right] \end{aligned} \quad (99)$$

Using the prior information that σ_j^2 is monotonically increasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1,\dots,j\}}^2)$ and $\underline{k}_{\{1,\dots,M_1^{j-1}\}}$ is monotonically decreasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1,\dots,j\}}^2)$, and (98)-(99), we arrive at $\frac{dk_{\{1,\dots,M_1^j\}}}{d\sigma_1^2} < 0$ and $\frac{d\sigma_j^2}{d\sigma_1^2} > 0$ over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1,\dots,j\}}^2)$. Thus, the monotonicity properties of $\underline{k}_{\{1,\dots,M_1^j\}}$ and σ_j^2 are proved.

The above induction argument implies that $\hat{\sigma}_{\{1\}}^2 > \hat{\sigma}_{\{1,2\}}^2 > \dots > \hat{\sigma}_{\{1,\dots,J-1\}}^2 > 0$. Thus when $\sigma_1^2 \in (0, \hat{\sigma}_{\{1,\dots,J-1\}}^2)$, it is guaranteed that (41) holds. The inequality (41) plays an important role in showing that the Gaussian test channel achieves the optimal sum rate.

Case 3 (monotonicity of \underline{k}_w (or $\underline{k}_{\{1,\dots,L\}}$): We list the prior conditions for the derivation. From the analysis in **Case 1** and **2**, it is clear that $\underline{k}_{\{1,\dots,M_1^{J-1}\}}$ is monotonically decreasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1,\dots,J-1\}}^2)$, and bounded away from 0 and $+\infty$, i.e. $A_{J-1} < \underline{k}_{\{1,\dots,M_1^{J-1}\}} < \left(\sum_{i=1}^{M_1^{J-1}} k_i^{-1} \right)^{-1}$. Also σ_J^2 is monotonically increasing over $\sigma_1^2 \in (0, \hat{\sigma}_{\{1,\dots,J-1\}}^2)$, and can take any positive real value.

We work with (39). In order that \underline{k}_0 is positive, we consider the inequality

$$\begin{aligned} \sigma_J^{-2} &> \left(\sigma_J^2 + \underline{k}_{\{1,\dots,M_1^{J-1}\}} \right)^{-1} + \sum_{i=1}^{m_J} \left(\sigma_J^2 + k_{M_1^{J-1}+i} \right)^{-1} \\ \stackrel{\sigma_J^2 > 0}{\Rightarrow} \sum_{i=1}^{m_J} \frac{k_{M_1^{J-1}+i}}{\sigma_J^2 + k_{M_1^{J-1}+i}} + \frac{\underline{k}_{\{1,\dots,M_1^{J-1}\}}}{\sigma_J^2 + \underline{k}_{\{1,\dots,M_1^{J-1}\}}} &> m_J. \end{aligned} \quad (100)$$

Again using the fact that $\underline{k}_{\{1,\dots,M_1^{J-1}\}}$ is bounded away from 0 and $+\infty$, and the intermediate

value theorem, we conclude that there exists $0 < \ddot{\sigma}_1^2 < \hat{\sigma}_{\{1, \dots, J-1\}}^2$ such that

$$0 < \sigma_1^2 < \ddot{\sigma}_1^2 \Rightarrow \underline{k}_w > 0 \quad (101)$$

$$\text{and} \quad \begin{cases} \underline{k}_w \rightarrow \left(\sum_{i=1}^L k_i^{-1} \right)^{-1} \text{ as } \sigma_1^2 \rightarrow 0 \\ \underline{k}_w \rightarrow 0 \text{ as } \sigma_1^2 \rightarrow \ddot{\sigma}_1^2 \end{cases}. \quad (102)$$

To show that \underline{k}_w is monotonically increasing over $\sigma_1^2 \in (0, \ddot{\sigma}_1^2)$, one can take the differentiation of (39) w.r.t. σ_1^2 and use the monotonicity properties of σ_j^2 and $\underline{k}_{\{1, \dots, M_1^{J-1}\}}$. The argument is straightforward. The proof is complete.

APPENDIX E

PROOF OF LEMMA 2.5

To prove that the corresponding \mathbf{K}_w is positive definite, it is sufficient to show that $|\mathbf{K}_{\{1, \dots, j\}}| > 0$, $\forall j = 1, \dots, L$. Before presenting the proof, we first show that $\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}}$, $j = 1, \dots, J$ are positive definite. Using the fact that $\mathbf{0} \prec \alpha_1 \mathbf{A}_1 \prec \alpha_2 \mathbf{A}_2(\mathbf{A}_1) \prec \dots \prec \alpha_J \mathbf{A}_J(\mathbf{A}_1)$ and (16), we arrive at

$$\underline{\mathbf{K}}_w \prec \underline{\mathbf{K}}_{\{1, \dots, M_1^{J-1}\}} \prec \dots \prec \underline{\mathbf{K}}_{\{1, \dots, M_1^1\}}. \quad (103)$$

We now use induction argument to prove the lemma. Assume $\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} \succ \mathbf{0}$, we show that the determinant of $\mathbf{K}_{\{1, \dots, M_1^{j-1}+i\}}$ is positive for any $i = 1, \dots, m_j$. For the special case that $j = 1$, we make no assumption. The argument is provided as follows:

$$\begin{aligned} & \left| \mathbf{K}_{\{1, \dots, M_1^{j-1}+i\}} \right| \\ &= \left| \text{diag}(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j, \mathbf{K}_{M_1^{j-1}+1} + \mathbf{A}_j, \dots, \mathbf{K}_{M_1^{j-1}+i} + \mathbf{A}_j) - (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \mathbf{A}_j (\mathbf{I}_N, \dots, \mathbf{I}_N) \right| \\ &= \left| \mathbf{I}_{N(M_1^{j-1}+i)} - (\mathbf{I}_N, \dots, \mathbf{I}_N)^t \mathbf{A}_j (\mathbf{I}_N, \dots, \mathbf{I}_N) \cdot \text{diag} \left(\left(\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right)^{-1}, \right. \right. \\ & \quad \left. \left. \left(\mathbf{K}_{M_1^{j-1}+1} + \mathbf{A}_j \right)^{-1}, \dots, \left(\mathbf{K}_{M_1^{j-1}+i} + \mathbf{A}_j \right)^{-1} \right) \right| \cdot \left| \mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right| \cdot \prod_{k=1}^i \left| \mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j \right| \\ &\stackrel{(a)}{=} \left| \mathbf{A}_j \right| \cdot \left| \mathbf{A}_j^{-1} - \left(\sum_{k=1}^i \left(\mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j \right)^{-1} + \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{A}_j \right)^{-1} \right) \right| \cdot \left| \mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j \right| \\ & \quad \cdot \prod_{k=1}^i \left| \mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j \right|, \end{aligned} \quad (104)$$

where (a) follows from (85) and the determinant identity $|\mathbf{I}_m - \mathbf{A}_{m \times n} \mathbf{B}_{n \times m}| = |\mathbf{I}_n - \mathbf{B}_{n \times m} \mathbf{A}_{m \times n}|$. Since $\mathbf{K}_{\{1, \dots, M_1^{j-1}\}}$ is assumed to be positive definite, it is straightforward that $|\mathbf{K}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{H} \otimes \mathbf{A}_j| > 0$. From the condition of the lemma, it is also known that $|\mathbf{A}_j| > 0$ and $|\mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j| > 0$, $k = 1, \dots, m_j$. Now we prove that the remaining quantity in (104) is positive. From (17), there is

$$\begin{aligned} & \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}} + \mathbf{A}_j \right)^{-1} = \sum_{k=1}^{m_j} \left(\mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j \right)^{-1} + \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{A}_j \right)^{-1} \\ \Rightarrow & \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^j\}} + \mathbf{A}_j \right)^{-1} \succeq \sum_{k=1}^i \left(\mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j \right)^{-1} + \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{A}_j \right)^{-1} \\ \stackrel{(a)}{\Rightarrow} & \mathbf{A}_j^{-1} \succ \sum_{k=1}^i \left(\mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j \right)^{-1} + \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{A}_j \right)^{-1}, \end{aligned}$$

where (a) follows from (103) and Lemma A.2. It is immediate that $|\mathbf{A}_j^{-1} - \sum_{k=1}^i \left(\mathbf{K}_{M_1^{j-1}+k} + \mathbf{A}_j \right)^{-1} + \left(\underline{\mathbf{K}}_{\{1, \dots, M_1^{j-1}\}} + \mathbf{A}_j \right)^{-1}| > 0$. Thus we have proven that the determinant of $\mathbf{K}_{\{1, \dots, M_1^{j-1}+i\}}$ is positive for any $i = 1, \dots, m_j$, $j = 1, \dots, J$. We conclude that the considered \mathbf{K}_w is positive definite. The proof is complete.

REFERENCES

- [1] A. A. E. Gamal and T. M. Cover, "Achievable Rates for Multiple Description Quantizers," *IEEE Trans. Inform. Th.*, vol. IT-28, pp. 851–857, 1982.
- [2] L. Ozarow, "On a Source-Coding Problem with Two Channels and Three Receivers," *Bell system Technical Journal*, vol. 59, pp. 1909–1921, 1980.
- [3] Z. Zhang and T. Berger, "New Results in Binary Multiple Descriptions," *IEEE. Trans. Inf. Theory*, vol. IT-33, no. 4, pp. 502–521, 1987.
- [4] R. Venkataramani, G. Kramer, and V. K. Goyal, "Multiple Description Coding with Many Channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 9, pp. 2106–2114, 2003.
- [5] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-Channel Symmetric Multiple Descriptions-part I: (n,k) Source-Channel Erasure Codes," *IEEE Trans. Inform. Th.*, vol. 50, pp. 47–61, 2004.
- [6] R. Puri, S. S. Pradhan, and K. Ramchandran, "n-Channel Symmetric Multiple Descriptions- part II: An Achievable Rate-Distortion Region," *IEEE Trans. Inf. Theory*, pp. 51:(4):1377–1392, 2005.
- [7] C. Tian, S. Mohajer, and S. N. Diggavi, "Approximating the Gaussian Multiple Description Rate Region Under Symmetric Distortion Constraints," *submitted to IEEE Trans. Inf. Theory*, 2008.
- [8] H. Wang and P. Viswanath, "Vector Gaussian Multiple Description With Individual and Central Receivers," *IEEE Trans. Inform. Theory*, vol. 53, pp. 2133–2153, 2007.

- [9] —, “Vector Gaussian Multiple Description with Two Levels of Receivers,” *IEEE Trans. Inform. Theory*, vol. 55, pp. 401–410, 2009.
- [10] J. Chen, “Rate Region of Gaussian Multiple Description Coding With Individual and Central Distortion Constraints,” *IEEE Trans. Inform. Theory*, vol. 55, pp. 3991–4005, 2009.
- [11] R. Zamir, “Gaussian Codes and Shannon Bounds for Multiple Descriptions,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 2629–2636, 1999.
- [12] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.
- [13] H. Weingarten, Y. Steinberg, and S. Shamai, “The Capacity Region of the Gaussian multiple-Input Multiple-Output Broadcast Channel,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 3936–3964, 2006.
- [14] D. J. A. Welsh, *Matroid Theory*. U.K.: Academic, 1976.
- [15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [16] F. Zhang, *Matrix Theory: Basic Results and Techniques*. Springer, 1999.
- [17] R. Liu, T. Liu, H. V. Poor, and S. Shamai, “A Vector Generalization of Costa’s Entropy-Power Inequality with Applications,” *submitted to IEEE Trans. Inf. Theory*, 2009.
- [18] G. Strang, *Linear Algebra and Its Applications*. Wellesley-Cambridge Press, 1998.
- [19] H. L. Royden, *Real Analysis*. NJ: Prentice-Hall, 1988.